MODULATORS OF BODY WEIGHT, CORRESPONDING NUCLEIC ACIDS AND PROTEINS, AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF

The research leading to the present inventions was funded in part by Grant No. DK 41096 from the National Institutes of Health. The government may have certain rights in the invention.

RELATED APPLICATIONS

The present application is a continuation-in-part of copending application Serial No. 08/347,563, filed November 30, 1994, which in turn is a continuation-in-part of copending application Serial No. 08/292,345, filed August 17, 1994, of which the instant application claims the benefit of the filing date pursuant to 35 U.S.C. § 120, and each of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to the control of body weight of mammals including animals and humans, and more particularly to materials identified herein as modulators of weight, and to the diagnostic and therapeutic uses to which such modulators may be put.

BACKGROUND OF THE INVENTION

Obesity, defined as an excess of body fat relative to lean body mass, is associated with important psychological and medical morbidities, the latter including hypertension, elevated blood lipids, and Type II or non-insulin-dependent diabetes melitis (NIDDM). There are 6-10 million individuals with NIDDM in the U.S., including 18% of the population of 65 years of age (Harris *et al.*, 1987). Approximately 45% of males and 70% of females with NIDDM are obese, and their diabetes is substantially improved or eliminated by weight reduction (Harris, 1991). As described below, both obesity and NIDDM are strongly heritable,

though the predisposing genes have not been identified. The molecular genetic basis of these metabolically related disorders is an important, poorly understood problem.

The assimilation, storage, and utilization of nutrient energy constitute a complex homeostatic system central to survival of metazoa. Among land-dwelling mammals, storage in adipose tissue of large quantities of metabolic fuel as triglycerides is crucial for surviving periods of food deprivation. The need to maintain a fixed level of energy stores without continual alterations in the size and shape of the organism requires the achievement of a balance between energy intake and expenditure. However, the molecular mechanisms that regulate energy balance remain to be elucidated. The isolation of molecules that transduce nutritional information and control energy balance will be critical to an understanding of the regulation of body weight in health and disease.

An individual's level of adiposity is, to a large extent, genetically determined. Examination of the concordance rates of body weight and adiposity amongst mono- and dizygous twins or adoptees and their biological parents have suggested that the heritability of obesity (0.4-0.8) exceeds that of many other traits commonly thought to have a substantial genetic component, such as schizophrenia, alcoholism, and atherosclerosis (Stunkard *et al.*, 1990). Familial similarities in rates of energy expenditure have also been reported (Bogardus *et al.*, 1986). Genetic analysis in geographically delimited populations has suggested that a relatively small number of genes may account for the 30%-50% of variance in body composition (Moll *et al.*, 1991). However, none of the genes responsible for obesity in the general population have been genetically mapped to a definite chromosomal location.

Rodent models of obesity include seven apparently single-gene mutations. The most intensively studied mouse obesity mutations are the *ob* (obese) and *db*

(diabetes) genes. When present on the same genetic strain background, *ob* and *db* result in indistinguishable metabolic and behavioral phenotypes, suggesting that these genes may function in the same physiologic pathway (Coleman, 1978). Mice homozygous for either mutation are hyperphagic and hypometabolic, leading to an obese phenotype that is notable at one month of age. The weight of these animals tends to stabilize at 60-70 g (compared with 30-35 g in control mice). *ob* and *db* animals manifest a myriad of other hormonal and metabolic changes that have made it difficult to identify the primary defect attributable to the mutation (Bray *et al.*, 1989).

Each of the rodent obesity models is accompanied by alterations in carbohydrate metabolism resembling those in Type II diabetes in man. In some cases, the severity of the diabetes depends in part on the background mouse strain (Leiter, 1989). For both *ob* and *db*, congenic C57BL/Ks mice develop a severe diabetes with ultimate β cell necrosis and islet atrophy, resulting in a relative insulinopenia. Conversely, congenic C57BL/6J *ob* and *db* mice develop a transient insulin-resistant diabetes that is eventually compensated by β cell hypertrophy resembling human Type II diabetes.

The phenotype of *ob* and *db* mice resembles human obesity in ways other than the development of diabetes - the mutant mice eat more and expend less energy than do lean controls (as do obese humans). This phenotype is also quite similar to that seen in animals with lesions of the ventromedial hypothalamus, which suggests that both mutations may interfere with the ability to properly integrate or respond to nutritional information within the central nervous system. Support for this hypothesis comes from the results of parabiosis experiments (Coleman, 1973) that suggest *ob* mice are deficient in a circulating satiety factor and that *db* mice are resistant to the effects of the ob factor (possibly due to an ob receptor defect). These experiments have led to the conclusion that obesity in these mutant mice may result from different defects in an afferent loop and/or integrative center of the postulated feedback mechanism that controls body composition.

Using molecular and classical genetic markers, the *ob* and *db* genes have been mapped to proximal chromosome 6 and midchromosome 4, respectively (Bahary *et al.*, 1990; Friedman *et al.*, 1991b). In both cases, the mutations map to regions of the mouse genome that are syntonic with human, suggesting that, if there are human homologs of *ob* and *db*, they are likely to map, respectively, to human chromosomes 7q and 1p. Defects in the *db* gene may result in obesity in other mammalian species: in genetic crosses between Zucker *fa/fa* rats and Brown Norway +/+ rats, the *fa* mutation (rat chromosome 5) is flanked by the same loci that flank *db* in mouse (Truett *et al.*, 1991).

Because of the myriad factors that seem to impact body weight, it is difficult to speculate as to which of these factors, and more particularly, which homeostatic mechanism is actually primarily determinative. Nonetheless, the apparent connection between the *ob* gene and the extent and characteristics of obesity have prompted the further investigation and elucidation that is reflected by the present application. It is the identification of the sequence of the gene and corresponding peptide materials, to which the present invention following below directs itself.

The citation of any reference herein should not be construed as an admission that such reference is prior art to the instant invention. Full citations of references cited by author and year are found at the end of the specification.

SUMMARY OF THE INVENTION

In its broadest aspect, the present invention relates to the elucidation and discovery of nucleotide sequences, and proteins putatively expressed by such nucleic acids or degenerate variations thereof, that demonstrate the ability to participate in the control of mammalian body weight. The nucleotide sequences in object are believed to represent the genes corresponding to the murine and human *ob* gene, that is postulated to play a critical role in the regulation of body weight and adiposity. Data presented herein indicates that the polypeptide product of the gene

in question is secreted by the cells that express it and that the polypeptide functions as a hormone.

In addition, the Examples herein demonstrate that the ob polypeptide, alternatively termed herein "leptin," circulates in mouse, rat, and human plasma. Leptin is absent in plasma from ob/ob mice, and is present at ten-fold higher concentrations in plasma from db/db mice, and twenty-fold higher concentrations in fa/fa rats. Most significantly, daily injections of recombinant leptin dramatically reduces the body mass of ob/ob mice, significantly effects the body weight of wild-type mice, and has no effect on db/db mice.

In a further aspect, the ob polypeptide from one species is biologically active in another species. In particular, the human ob polypeptide is active in mice.

In a first instance, the modulators of the present invention comprise nucleic acid molecules, including recombinant DNA molecules (*e.g.*, cDNA or a vector containing the cDNA or isolated genomic DNA) or cloned genes (*i.e.* isolated genomic DNA), or degenerate variants thereof, which encode polypeptides themselves serving as modulators of weight control as hereinafter defined, or conserved variants or fragments thereof, particularly such fragments lacking the signal peptide (alternatively referred to herein as mature ob polypeptide), which polypeptides possess amino acid sequences such as set forth in FIGURE 1 (SEQ ID NO:2), FIGURE 3 (SEQ ID NO:4), FIGURE 5 (SEQ ID NO:5) and FIGURE 6 (SEQ ID NO:6). In specific embodiments, amino acid sequences for two variants of murine and human ob polypeptides are provided. Both polypeptides are found in a form with glutamine 49 deleted, which may result from an mRNA splicing anomaly. The ob polypeptides from various species may be highly homologous; as shown in Figure 4, murine and human ob polypeptides are greater than 80% homologous.

The nucleic acid molecules, recombinant DNA molecules, or cloned genes, may have the nucleotide sequences or may be complementary to DNA coding sequences shown in FIGURE 1 (SEQ ID NO:1) and FIGURE 2 (SEQ ID NO:3). In particular, such DNA molecules can be cDNA or genomic DNA isolated from the chromosome. Nucleic acid molecules of the invention may also correspond to 5′ and 3′ flanking sequences of the DNA. Accordingly, the present invention also relates to the identification of a gene having a nucleotide sequence selected from the sequences of Figure 1 (SEQ ID NO:1) and Figure 2 (SEQ ID NO:3) herein, and degenerate variants, allelic variations, and like cognate molecules.

A nucleic acid molecule of the invention can be DNA or RNA, including synthetic variants thereof having phosphate or phosphate analog, *e.g.*, thiophosphate, bonds. Both single stranded and double stranded sequences are contemplated herein.

The present invention further provides nucleic acid molecules for use as molecular probes, or as primers for polymerase chain reaction (PCR) amplification, *i.e.*, synthetic or natural oligonucleotides having a sequence corresponding to a portion of the sequences shown in Figure 1 (SEQ ID NO:1), Figure 2 (SEQ ID NO:3) and Figure 20A (SEQ ID NO:22), or the 5′ and 3′ flanking sequences of the coding sequences. In particular, the invention contemplates a nucleic acid molecule having at least about 10 nucleotides, wherein a sequence of the nucleic acid molecule corresponds to a nucleotide sequence of the same number of nucleotides in the nucleotide sequences of Figure 1 (SEQ ID NO:1), Figure 2 (SEQ ID NO:3) and Figure 20A (SEQ ID NO:22), or a sequence complementary thereto. More preferably, the nucleic acid sequence of the molecule has at least 15 nucleotides. Most preferably, the nucleic acid sequence has at least 20 nucleotides. In an embodiment of the invention in which the oligonucleotide is a probe, the oligonucleotide is detectably labeled, *e.g.*, with a radionuclide (such as ^{32}P), or an enzyme.

In further aspects, the present invention provides a cloning vector, which comprises the nucleic acids of the invention that encode the ob polypeptide; and a bacterial, insect, or a mammalian expression vector, which comprises the nucleic acid molecules of the invention encoding the ob polypeptide, operatively associated with an expression control sequence. Accordingly, the invention further relates to a host cell, such as a bacterial cell, yeast cell, insect cell, or a mammalian cell, transfected or transformed with an appropriate expression vector, and correspondingly, to the use of the above mentioned constructs in the preparation of the modulators of the invention.

In yet a further aspect, the present invention relates to antibodies that bind to the ob polypeptide. Such antibodies may be generated against the full length polypeptide, or antigenic fragments thereof. In one aspect, such antibodies inhibit the functional (*i.e.*, body weight and fat composition modulating) activity of the ob polypeptide. In another aspect, antibodies can be used to determine the level of circulating ob polypeptide in plasma or serum. In yet a further aspect, regio-specific antibodies, particularly monoclonal antibodies, can be used as probes of ob polypeptide structure.

All of the foregoing materials are to be considered herein as modulators of body weight and fat composition, and as such, may be used in a variety of contexts. Specifically, the invention contemplates both diagnostic and therapeutic applications, as well as certain agricultural applications, all contingent upon the use of the modulators defined herein, including both nucleic acid molecules and peptides. Moreover, the modulation of body weight carries specific therapeutic implications and benefits, in that conditions where either obesity or, conversely, cachexia represent undesired bodily conditions, can be remedied by the administration of one or more of the modulators of the present invention.

Thus, a method for modulating body weight of a mammal is proposed that comprises controlling the expression of the protein encoded by a nucleic acid

having nucleotide sequence selected from the sequence of Figure 1 (SEQ ID NO:1), the sequence of Figure 2 (SEQ ID NO:3) and degenerate and allelic variants thereof. Such control may be effected by the introduction of the nucleotides in question by gene therapy into fat cells of the patient or host to control or reduce obesity. Conversely, the preparation and administration of antagonists to the nucleotides, such as anti-sense molecules, would be indicated and pursued in the instance where conditions involving excessive weight loss, such as anorexia nervosa, cancer, or AIDS are present and under treatment. Such constructs would be introduced in similar fashion to the nucleotides, directly into fat cells to effect such changes.

Correspondingly, the proteins defined by Figures 1, 3, 5, and 6 (SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6), conserved variants, active fragments thereof, and cognate small molecules could be formulated for direct administration for therapeutic purposes, to effect reduction or control of excessive body fat or weight gain. Correspondingly, antibodies and other antagonists to the stated protein materials, such as fragments thereof, could be prepared and similarly administered to achieve the converse effect. Accordingly, the invention is advantageously directed to a pharmaceutical composition comprising an ob polypeptide of the invention, or alternatively an antagonist thereof, in an admixture with a pharmaceutically acceptable carrier or excipient.

The diagnostic uses of the present nucleotides and corresponding peptides extend to the use of the nucleotides to identify further mutations of allelic variations thereof, so as to develop a repertoire of active nucleotide materials useful in both diagnostic and therapeutic applications. In particular, both homozygous and heterozygous mutations of the nucleotides in question could be prepared that would be postulated to more precisely quantitate the condition of patients, to determine the at-risk potential of individuals with regard to obesity. Specifically, heterozygous mutations are presently viewed as associated with mild to moderate obesity, while homozygous mutations would be associated with a more pronounced

and severe obese condition. Corresponding DNA testing could then be conducted utilizing the aforementioned ascertained materials as benchmarks, to facilitate an accurate long term prognosis for particular tendencies, so as to be able to prescribe changes in either dietary or other personal habits, or direct therapeutic intervention, to avert such conditions.

The diagnostic utility of the present invention extends to methods for measuring the presence and extent of the modulators of the invention in cellular samples or biological extracts (or samples) taken from test subjects, so that both the encoded nucleotide (genomic DNA or RNA) and or the levels of protein in such test samples could be ascertained. Given that the increased activity of the nucleotide and presence of the resulting protein reflect the capability of the subject to inhibit obesity, the physician reviewing such results in an obese subject would determine that a factor other than dysfunction with respect to the presence and activity of the nucleotides of the present invention is a cause of the obese condition. Conversely, depressed levels of the nucleotide and or the expressed protein would suggest that such levels must be increased to treat such obese condition, and an appropriate therapeutic regimen could then be implemented.

Further, the nucleotides discovered and presented in Figures 1 and 2 represent cDNA in which, as stated briefly above, is useful in the measurement of corresponding RNA. Likewise, recombinant protein material corresponding to the polypeptides of Figures 1 and 3 may be prepared and appropriately labeled, for use, for example, in radioimmunoassays, for example, for the purpose of measuring fat and/or plasma levels of the ob protein, or for detecting the presence and level of a receptor for ob on tissues, such as the hypothalamus.

Yet further, the present invention contemplates not only the identification of the nucleotides and corresponding proteins presented herein, but the elucidation of the receptor to such materials. In such context, the polypeptides of Figures 1, 3, 5, and/or 6 could be prepared and utilized to screen an appropriate expression library

to isolate active receptors. The receptor could thereafter be cloned, and the receptor alone or in conjunction with the ligand could thereafter be utilized to screen for small molecules that may possess like activity to the modulators herein.

Yet further, the present invention relates to pharmaceutical compositions that include certain of the modulators hereof, preferably the polypeptides whose sequences are presented in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:6, their antibodies, corresponding small molecule agonists or antagonists thereof, or active fragments prepared in formulations for a variety of modes of administration, where such therapy is appropriate. Such formulations would include pharmaceutically acceptable carriers, or other adjuvants as needed, and would be prepared in effective dosage ranges to be determined by the clinician or the physician in each instance.

Accordingly, it is a principal object of the present invention to provide modulators of body weight as defined herein in purified form, that exhibit certain characteristics and activities associated with control and variation of adiposity and fat content of mammals.

It is a further object of the present invention to provide methods for the detection and measurement of the modulators of weight control as set forth herein, as a means of the effective diagnosis and monitoring of pathological conditions wherein the variation in level of such modulators is or may be a characterizing feature.

It is a still further object of the present invention to provide a method and associated assay system for the screening of substances, such as drugs, agents and the like, that are potentially effective to either mimic or inhibit the activity of the modulators of the invention in mammals.

It is a still further object of the present invention to provide a method for the treatment of mammals to control body weight and fat content in mammals, and or

to treat certain of the pathological conditions of which abnormal depression or elevation of body weight is a characterizing feature.

It is a still further object of the present invention to prepare genetic constructs for use in genetic therapeutic protocols and or pharmaceutical compositions for comparable therapeutic methods, which comprise or are based upon one or more of the modulators, binding partners, or agents that may control their production, or that may mimic or antagonize their activities.

Other objects and advantages will become apparent to those skilled in the art from a review of the ensuing description which proceeds with reference to the following illustrative drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 depicts the nucleic acid sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) derived for the murine *ob* cDNA. A 97 base pair 5′ leader was followed by a predicted 167 amino acid open reading frame and an approximately 3700 kB 3′ untranslated sequence. A total of about 2500 base pairs of the 3′ unstranslated sequence is shown. Analysis of the predicted protein sequence by observation and using the *SigSeq* computer program indicates the presence of a signal sequence (underlined). Microheterogeneity of the cDNA was noted in that approximately 70% of the cDNAs had a glutamine codon at codon 49 and 30% did not (see FIGURES 5 and 6, *infra*). This amino acid is underlined, as is the arginine codon that is mutated in C57BL/6J ob/ob mice (1J mice).

FIGURE 2 depicts the nucleic acid sequence (SEQ ID NO:3) derived for the human *ob* cDNA. The nucleotides are numbered from 1 to 701 with a start site at nucleotide 46 and a termination at nucleotide 550.

FIGURE 3 depicts the full deduced amino acid sequence (SEQ ID NO:4) derived for the human *ob* gene corresponding to the nucleic acid sequence of FIGURE 2. The amino acids are numbered from 1 to 167. A signal sequence cleavage site is located after amino acid 21 (Ala) so that the mature protein extends from amino acid 22 (Val) to amino acid 167 (Cys).

FIGURE 4 depicts the comparison between the murine (SEQ ID NO:2) and human (SEQ ID NO:4) deduced amino acid sequences. The sequence of the human *ob* deduced amino acid sequence was highly homologous to that of mouse. Conservative changes are noted by a dash, and non-conservative changes by an asterisk. The variable glutamine codon is underlined, as is the position of the nonsense mutation in C57BL/6J ob/ob (1J) mice. Overall, there is 84% identity at the amino acid level, although only six substitutions were found between the valine at codon 22 (immediately downstream of the signal sequence overage) and the cysteine at position 117.

FIGURE 5 depicts the full length amino acid sequence (SEQ ID NO:5) derived for the murine *ob* gene as shown in FIGURE 3, but lacking glutamine at position 49. The nucleotides are numbered from 1 to 166. A signal sequence cleavage site is located after amino acid 21 (Ala) (and thus, before the glutamine 49 deletion) so that the mature protein extends from amino acid 22 (Val) to amino acid 166 (Cys).

FIGURE 6 depicts the full deduced amino acid sequence (SEQ ID NO:6) derived for the human *ob* gene as shown in FIGURE 4, but lacking glutamine at position 49. The nucleotides are numbered from 1 to 166. A signal sequence cleavage site is located after amino acid 21 (Ala) (and thus, before the glutamine 49 deletion) so that the mature protein extends from amino acid 22 (Val) to amino acid 166 (Cys).

FIGURE 7. (A) Physical map of the location of *ob* in the murine chromosome, and the YAC and P1 cloning maps. "M and N" corresponds to *Mul*I and *Not*I restriction sites. The numbers correspond to individual animals that were

recombinant in the region of *ob* of the 1606 meioses that were scored. Met, Pax 4, D6Rck39, D6Rck13, and Cpa refer to locations in the region of *ob* that bind to the DNA probes. YACs were isolated using D6Rck13 and Pax-4 as probes, and the ends were recovered using vectorette PCR and/or plasmid end rescue and used in turn to isolate new YACs. (B) The resulting YAC contig. One of the YACs in this contig, Y902A0925, was chimeric. Each of the probes used to genotype the recombinant animals is indicated in parentheses. (6) Corresponds to YAC 107; (5) corresponds to M16(+) (or M16(pLUS)); (4) corresponds to *adu*(+); (3) corresponds to *aad*(pICL); (2) corresponds to 53(pICL); and (1) corresponds to 53(+). (C) The P1 contig of bacteriophage P1 clones isolated with selected YAC end probes. The *ob* gene was isolated in a P1 clone isolated using the distal end of YAC YB6S2F12 (end (4)) (alternatively termed herein *adu*(+)).

FIGURE 8 presents a photograph of an ethidium bromide stain of 192 independent isolates of the fourth exon trapping experiment that were PCR amplified and characterized.

FIGURE 9 is a photograph of an ethidium bromide stain of PCR-amplified clones suspected of carrying *ob*. Each of the 7 clones that did not carry the artifact was reamplified using PCR and electrophoresed on a 1% agarose gel in TBE and stained with ethidium bromide. The size markers (far left unnumbered lane) are the commercially available "1 kB ladder". Lane 1 -- clone 1D12, containing an "HIV sequence." Lane 2 -- clone 1F1, a novel clone outside of the *ob* region. Lane 3 -- clone 1H3. Lane 4 -- clone 2B2, which is the identical to 1F1. Lane 5 -- clone 2G7, which contains an *ob* exon. Lane 6 -- clone 2G11, which is identical to 1F1. Lane 7 -- clone 2H1, which does not contain an insert.

FIGURE 10 presents the sequence of the 2G7 clone (SEQ ID NO:7), which includes an exon coding for a part of the *ob* gene. The primer sequences used to amplify this exon are boxed in the figure (SEQ ID NOS:8 and 9).

FIGURE 11. (A) Reverse transcription-PCR analysis of mRNA from different tissues of the same mouse with the 2G7 primers and actin primers. The RT-PCR reactions were performed using 100 ng of total RNA reverse transcribed with oligo dT as a primer for first strand cDNA. PCR amplification was performed for 35 cycles with 94° denaturation x 1'; 55° hybridization x 1'; and 72 extensions for 2' with a 1' second autoextension per cycle. RT-PCR products were resolved in a 2% low melting point agarose gel run in 1x TBE buffer. (B) Northern blot of mRNA from different organs of the mouse using PCR labeled 2G7 as a probe. Ten μg of total RNA from each of the tissues was electrophoresed on an agarose gel with formaldehyde. The probe was hybridized at 65°C in Rapid Hybe (Amersham). Autoradiographic signals were apparent after 1 hour of exposure; the experiment shown was the result of a 24 hour exposure.

FIGURE 12. (A) An ethidium bromide stain from an RT PCR reaction on fat cell (white adipose tissue) RNA from each of the mouse strains listed. Total RNA (100 ng) for each sample was reverse transcribed using oligo dT and reverse transcriptase, and the resulting single stranded cDNA was PCR amplified with the 2G7 primers (lower bands) or actin primers (upper bands). Both the 2G7 and actin primers were included in the same PCR reaction. The products were run on a 1% agarose TBE gel. (B) Northern analysis corresponding to (A). Ten μg of fat cell (white adipose tissue) RNA from each of the strains indicated were run out and probed with the PCR labeled 2G7 probe as in Figure 11B, above. An approximately 20-fold increase in the level of 2G7 mRNA was apparent in white fate RNA from the C57BL/6J ob/ob (1J) strain relative to lean littermates. In both the RT-PCR and Northern experiments there was no detectable signal in 2G7 RNA from the SM/Ckc-+^{Dac}ob^{2J}/ob^{2J} (2J) mice even after a 2 week exposure. A 24 hour autoradiographic exposure is shown. The same filter was hybridized to an actin probe (bottom portion of the panel).

FIGURE 13 is a Northern analysis of additional 2J animals and control animals that confirms the absence of the *ob* mRNA from 2J animals. The Northern

analysis was performed as in Figures 11 and 12. In this case, the control RNA was ap2, a fat specific transcript. There is no significance to the varying density of the ap2 bands.

FIGURE 14 compares the DNA sequence of the C57BL/6J (normal) and the C57BL/6J ob/ob (1J) mice in the region of the point mutation that leads to introduction of a premature stop codon (nonsense mutation) in the mutant strain cDNA. The ob/ob mice had a C→T mutation that changed an arginine residue at position 105. This base change is shown as the output from the automated DNA sequencer. RT-PCR was performed using white fat RNA from both strains (+/+ and ob/ob) using primers from the 5′ and 3′ unstranslated regions. The PCR reaction products were gel purified and directly sequenced manually and using an ABI 373A automated sequencer with primers along both strands of the coding sequence.

FIGURE 15. (A) Genomic southern blot of genomic DNA from each of the mouse strains listed. Approximately 5 μg of DNA (derived from genomic DNA prepared from liver, kidney or spleen) was restriction digested with the restriction enzyme indicated. The DNA was then electrophoresed in a 1% agarose TBE gel and probed with PCR labeled 2G7. Restriction digestion with *Bgl*II revealed an increase in the size of an approximately 9 kB (the largest) *Bgl*II fragment in SM/Ckc-+Dacob^{2J}/ob^{2J} (2J) DNA. RFLPs were not detectable with any other restriction enzymes. Preliminary restriction mapping of genomic DNA indicated that the polymorphic *Bgl*II site is about 7 kB upstream of the transcription start site. None of the other enzymes tested extend past the mRNA start site. (B) Segregation of a *Bgl*II polymorphism in the SM/Ckc-+Dacob^{2J}/ob^{2J} strain. Six obese and five lean progeny from the same generation of the coisogenic SM/Ckc-+^{Dac}ob^{2J}/ob^{2J} (2J) colony were genotyped by scoring the *Bgl*II polymorphism as shown in (A). All of the phenotypically obese animals were homozygous for the larger allele of the polymorphic *Bgl* fragment. The DNA in

the "control" lane was prepared from an unrelated SM/Ckc-+^{Dac}+/+ mouse, bred separately from the SM/Ckc-+^{Dac}ob^{2J}/ob^{2J} colony.

FIGURE 16 is a Southern blot of EcoRI digested genomic DNA from the species listed, using an *ob* cDNA as a probe (*i.e.*, a zoo blot). Hybridization signals were detectable in every vertebrate sample, even after a moderate stringency hybridization. The cat DNA in this experiment was slightly degraded. The restricted DNA was run on a 1% agarose TBE gel, and transferred to an imobilon membrane for probing. The filter was hybridized at 65°C and washed in 2X SSC/0.2% SDS at 65°C twice for twenty minutes and exposed for 3 days using Kodak X-OMAT film.

FIGURE 17 presents the expression cloning region of vector pET-15b (Novagen).

FIGURE 18 presents analysis of the eluate from a His-binding resin (Ni) column for a recombinant mature murine ob fusion to a His-tag (A) and mature human ob fusion to a His-tag (B). Bacteria transformed with vectors pETM9 and pETH14, respectively. Upon induction with 1 mM IPTG at optimal conditions, the transformed bacteria were able to produce 100-300 μg/ml of ob fusion protein, primarily in the inclusion body. The inclusion body was solubilized with 6M guanidine-HCl or urea, and fusion protein (present in the lysis supernatant) was loaded on the His-binding resin (Ni) column in 10 ml of 1x binding buffer with urea. The column was eluted stepwise with 5 ml aliquots of 20 μM, 60 μM, and 300 μM imidazole, and finally with strip buffer. The aliquots were analyzed for the presence of ob polypeptide fusion on a 15% acrylamide gel. Each lane contains the equivalent of 100 μl of bacterial extract.

FIGURE 19. (A) *In vitro* translation of *ob* RNA. A human *ob* cDNA was subcloned into the pGEM vector. The plasmid was linearized and plus strand RNA was synthesized using sp6 polymerase. The *in vitro* synthesized RNA was translated in the presence or absence of canine pancreatic microsomal membranes.

An approximately 18 kD primary translation product was seen after *in vitro* translation. The addition of microsomal membranes to the reaction led to the appearance of a second translation product about 2 kD smaller than the primary translation product. The size of the translation product of interleukin-1α RNA, which lacks an encoded signal sequence, was unchanged by the addition of microsomal membranes. These data indicated the presence of a functional signal sequence. (B) *In vitro* translation in the presence or absence of proteinase K. Protease treatment resulted in complete proteolysis of the 18 kD primary translation product, while the 16 kD processed form was unaffected. Permeabilization of the microsome with 0.1% TRITON-X100 rendered the processed from protease sensitive. These results indicate that the product had translated into the lumen of the microsome.

FIGURE 20. (A) The sequence of the human *ob* gene (SEQ ID NO:22). (B) A schematic diagram of the murine *ob* gene. (C) A schematic diagram of the human *ob* gene. In both (B) and (C), the start and stop codons are underlined. There is no evidence of a first intron homologous to the mouse first intron in the human gene, but its existence cannot be excluded.

FIGURE 21 presents a schematic drawing of one of the cloning strategies employed to achieve recombinant expression of *ob* in pichia yeast. (A) Expression vector of *ob* with an α-mating factor signal sequence. (B) Schematic drawing of the structure of the recombinant fusion protein, including the amino acid sequence (SEQ ID NO:23) showing the *Xho*I site and putative KEX-2 and STE-13 cleavage sites, and the N-terminal surplus amino acids present after KEX-2 cleavage (SEQ ID NO:24). (C) An alternative strategy for producing mature ob under involves preparing a construct with an amino acid sequence corresponding to a *Xho*I cleavage site and a KEX-2 cleavage site immediately upstream of the mature ob polypeptide sequence (SEQ ID NO:25).

FIGURE 22 Alternative expression strategy in pichia. (A) Expression vector of an *ob* fusion with a His tag adopted from the pET expression system under control of the α-mating factor signal sequence. (B) Schematic drawing of the structure of the recombinant ob fusion protein containing a His tag, which includes the α-mating factor signal sequence, putative KEX-2 and STE-13 cleavage sites, the His-tag, and a thrombin cleavage site, and which would yield ob with three surplus N-terminal amino acid residues.

FIGURE 23. (A) PAGE analysis of expression of murine *ob* (both the microheterogenous forms, *i.e.*, containing and missing Gln 49) in transformed pichia yeast. The expected band of approximately 16 kD is visible in the transformed yeast culture fluid (second and third lanes), but not in culture fluid from non-transformed yeast (first lane). (B) PAGE analysis of partially purified recombinant ob polypeptide on carboxymethyl cellulose, a weak cation exchanger. A band of about 16 kD is very visible in fractions 3 and 4 from the column, which was eluted with 250 mM NaCl. Lane 1 -- loaded sample; lane 2 -- flow through; lanes 3-5 -- fractions eluted with 250 mM NaCl.

THE REMAINDER OF THIS PAGE IS INTENTIONALLY LEFT BLANK

Figure 24. The *ob* Protein Circulates in Mouse Plasma

24 A. Immunoprecipitations from Mouse Blood

.5 ml of mouse plasma was pre-cleared with unconjugated sepharose and incubated overnight with immunopurified anti-*ob* antibodies conjugated to sepharose 4B beads. The immunoprecipitate was separated on a 15% SDS-PAGE gel, transferred and Western blotted with an anti-*ob* antibody. The protein migrated with a molecular weight of ~16 kD, to the same position as the mature mouse *ob* protein expressed in yeast. The protein was absent in plasma from C57BL/6J ob/ob mice and increased ten-fold in plasma from C57BLB/Ks db/db mice relative to wild type mice. *db* mice have been suggested to overproduce the *ob* protein, secondary to resistance to its effects.

24 B. Increased Levels of *ob* in fatty rats

The fatty rat is obese as a result of a recessive mutation on rat chromosome 5. Genetic data has suggested a defect in the same gene as is mutant in *db* mice. Plasma from fatty rats and lean littermates was immunoprecipitated and run on Western blots. A twenty-fold increase in the circulating level of *ob* is seen in the mutant animals.

24 C. Quantitation of the *ob* Protein in Mouse Plasma

Increasing amounts of the recombinant mouse protein were added to 100 λ of plasma from *ob* mice and immunoprecipitated. The signal intensity on Western blots was compared to that from 100 λ of plasma from wild type mice. A linear increase in signal intensity was seen with increasing amounts of recombinant protein demonstrating that the immunoprecipitations were performed under conditions of antibody excess. Similar signals were seen in the wild type plasma sample and the sample with 2 ng of recombinant protein indicating the circulating level in mouse plasma is ~20 ng/ml.

24 D. *ob* Protein in Adipose Tissue Extracts

Cytoplasmic extracts of mouse adipose tissue were prepared from *db* and wild type mice. Western blots showed increased levels of the 16 kD protein in extracts prepared from *db* mice.

Figure 25 The *ob* Protein Circulates at Variable Levels in Human Plasma

25A. Western Blots of Human Plasma

Plasma samples were obtained from six lean volunteers. Immunoprecipitation and Western blotting revealed the presence of an immunoreactive 16 kD protein, identical in size to a recombinant 146 amino acid human protein expressed in yeast. Variable levels of the protein were seen in each of the six samples.

25B. An ELISA (Enzyme Linked Immunoassay) for Human *ob*

Microtiter plates were coated with immunopurified anti-human *ob* antibodies. Known amounts of recombinant protein were added to the plates and detected using immunopurified biotinylated anti-*ob* antibodies. The resulting standard curve showed that the assay was capable of detecting 1 ng/ml or more of the human *ob* protein.

25C. Quantitation of the *ob* Protein in Human Plasma

An ELISA immunoassay was performed using 100 λ of plasma from the six lean volunteers and the standards used in panel B. Levels of the *ob* protein ranging from 2 ng/ml in HP1 to 15 ng/ml in HP6 were seen. These data correlated with the Western blot data in panel A.

Figure 26 The *ob* Protein Forms Inter- or Intramolecular Disulphide Bonds

26A. Western blots Under Non Reducing Conditions

The Western blots of mouse and human plasma were repeated with and without the addition of reducing agents to the sample buffer. When β-Mercaptoethanol is omitted from the sample buffer, immunoprecipitates from *db* plasma migrate with an apparent molecular mass of 16 kD and 32 kD. Addition of β-Mercaptoethanol to the buffer leads to the disappearance of the 32 kD moiety (see Fig. 1). This result is recapitulated when the mouse protein is expressed in the yeast, Pichia pastoris. In this case, the mouse *ob* protein migrates to the position of a dimer. Under reducing conditions the purified

recombinant mouse protein migrates with an apparent molecular weight of 16 kD indicating that the 32 kD molecular form is the result of one or two intermolecular disuphide bonds. The human protein expressed in vivo and in Pichia pastoris migrates with a molecular mass of 16 kD under both conditions (data not shown).

26B. The Human Protein Expressed in Yeast Contains an Intramolecular Disulphide Bond

Secreted proteins generally assume their correct conformation when expressed in the Pichia pastoris expression system. The 146 amino acid mature human protein was expressed in Pichia pastoris and purified from the yeast media by a two-step purification protocol involving IMAC and gel filtration. The purified recombinant protein was subjected to mass spectrometry before and after cyanogen bromide cleavage. Cyanogen bromide cleaves at the carboxy terminus of methionine residues. The molecular mass of the recombinant yeast protein was 16,024±3 Da (calculated molecular mass = 16,024 Da). Cyanogen bromide cleaves after the three methionines in the protein sequence at amino acids 75, 89 and 157. The cyanogen bromide fragment with measured mass 8435.6 Da corresponds to amino acids 90-157 and 158-167 joined by a disulphide linkage between cys-117 and cys-167 (calculated molecular mass = 8434.5 Da).

Figure 27 Preparation of Bioactive Recombinant Protein

The nucleotide sequence corresponding to the 145 amino acid mature mouse *ob* ptotein was cloned into the PET 15b expression vector. This PET vector inserts a polyhistidine tract (His-tag) upstream of the cloned sequence which allows efficient purification using Immobilized Metal Affinity Chromatography (IMAC). The recombinant bacterial protein initially partitioned in the insoluble membrane fraction after bacterial lysis. The membrane fraction was solubilized using guanidium hydrochloride and loaded onto an IMAC column. The protein was eluted stepwise with increasing concentrations of imidazole as shown. The eluted protein was refolded and treated with thrombin to remove the His-tag, as described below. The final yield of soluble protein was 45 ng/ml of bacterial culture.

Figure 28 Biologic Effects of the *ob* Protein

Fig. 28A. Time Course of Food Intake and Body Weight

Groups of ten animals received either daily intraperitoneal injections of the *ob* protein at a dose of 5 µg/kg/day, daily injections of PBS or no treatment. The treatment groups included C57Bl/6J ob/ob mice (left panels), C57Bl/Ks db/db mice (center panels) and CBA/J+/+ mice (right panels). The food intake of the mice was measured daily and the body weight was recorded at three to four day intervals as indicated. (The scale of the body weight in grams is different for the wild type mice vs. the *ob* and *db* mice.) The food intake of the *ob* mice receiving protein was reduced after the first injection and stabilized after the fourth day at a level ~40% of that seen in the sham injected group ($p < .001$). The body weight of these animals decreased an average of 1.3 grams/day and stabilized after three weeks to a level ~60% of the starting weight ($p < .0001$). No effect of the protein was demonstrable in *db* mice. Small but significant effects on body weight were observed in CBA/J mice at two early time points ($p < .02$). The standard error of each measure is depicted by a bar and the statistical significance of these results is shown in Table 1.

Fig. 28B. Pair Feeding of *ob* Mice

A group of four C57Bl/6J ob/ob mice were fed an amount of food equal to that consumed by the group of *ob* mice receiving recombinant protein. The weight loss for both groups was calculated after five, eight and twelve days. The food restricted mice loss less weight than the *ob* mice receiving protein ($p < .02$). This result indicates that the weight reducing effect of the *ob* protein is the result of effects on both food intake and energy expenditure.

Fig. 28C. Photograph of a Treated *ob* Mouse

Shown are two C57Bl/6J ob/ob mice. The mouse on the left received PBS and weighed 65 grams which was the starting weight. The mouse on the right received daily injections of the recombinant *ob* protein. The starting weight of this animal was also 65 grams, and the weight after three weeks of protein treatment was 38 grams.

Fig. 28. Livers From Treated and Untreated *ob* Mice

Shown are livers from treated and untreated C57Bl/6J ob/ob mice. The liver from the mouse receiving PBS had the gross appearance of a fatty liver and weighed 5.04 grams. The liver from the mouse receiving the recombinant *ob* protein had a normal appearance and weighed 2.23 grams.

Figure 29 InSitu Hybridization of *ob* to Adipose Tissue

Sense and Antisense *ob* RNA was labeled in vitro using Sp6 and t7 polymerase and digoxigenin. The labeled RNAs were hybridized to paraffin embedded sections of adipose tissue from epididymal fat pads of eight week old C57Bl/Ks mice (labelled wild type) and C57Bl/Ks db/db mice (labelled db). In the figure, the lipid droplets appear as unstained vacuoles within cells. The cytoplasm is a thin rim at the periphery of the cells and is indistinguishable from the cell membrane X 65. Hybridization to all the adipocytes in the field was detected in the wild type sections only using the antisense probe and greatly increased levels were seen in the tissue sections from the db/db animals.

Figure 30 ob RNA Is Expressed in Adipocytes *in vivo* and *in vitro*

Total RNA (10 micrograms) from several different sources was electophoresed on Northern blots and hybridized to an *ob* probe. Firstly, differences in cell buoyancy after collagenase digestion was used to purify adipocytes. *ob* RNA was present only in the adipocyte fraction. (lane SV indicates the stromovascular fraction and A indicates the adipocyte fraction) in addition, *ob* RNA was not expressed in the undifferentiated 3T3-442 preadipocyte cells. (labelled U) Differentiated adipocytes from these cell lines expressed clearly detectable levels of *ob* mRNA (labelled D).

Figure 31 ob RNA is Expressed in All Adipose Tissue Depots

All of the adipose tissue depots tested expressed *ob* RNA. The inguinal fat pad expressed somewhat lower RNA levels although there was variability in the levels of signals in different experiments. (Fig.31A) Lanes 1) epididymal 2) inguinal 3)abdominal 4) parametrial fat pads. Brown fat also expressed a low level of *ob* RNA. (Fig.31B) The level of *ob* expression in brown fat was unchanged in animals housed at 4°C for one week while the abundance of the brown fat specific UCP RNA, known to be cold inducible, increased five-fold.

Figure 32 Expression of ob RNA in db/db and Gold ThioGlucose Treated Mice

Total RNA from the parametrial fat pads of db/db and Gold Thioglucose (GTG) treated mice was electrophoresed on a Northern blot. GTG administered as a single dose is known to cause obesity by inducing specific hypothalamic lesions. One month old CBA female mice were treated with GTG (.2 mg/g) with a resulting increase of >20 g in treated animals relative to control animals (<5 g). Hybridization of an *ob* probe to RNA from db/db and GTG treated mice revealed a twenty-fold increase in the abundance of *ob* RNA relative to control RNA (actin or GAPDH).

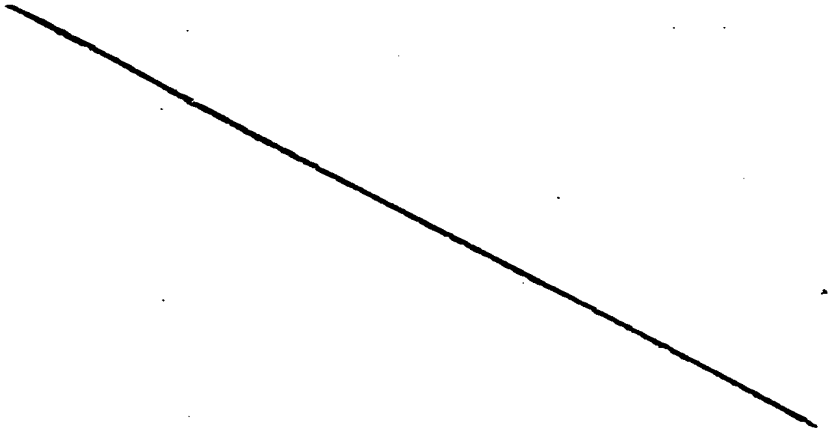

Figure 33 Northern blot analysis of human RNA. Northern blots containing 10 μg of total RNA from human adipose tissue (FAT, panel **A**) and 2 μg of polyA+ RNA from other human tissues (panel **B**) were hybridized to human OB or human β actin probes as indicated. An intense signal at ~4.5 kb was seen with the adipose tissue total RNA. Hybridization to the polyA+ RNA revealed detectable signals in heart (HE) and placenta (PL), whereas OB RNA was not detected in brain (BR), lung (LU), liver (LI), skeletal muscle (SM), kidney (KI), and pancreas (PA). In each case, the length of the autoradiographic exposure is indicated. Of note, the genesis of the lower molecular bands seen in placental RNA (e.g., alternate splicing, RNA degradation) is not known.

Figure 34 YAC contig containing the human OB gene and 8 microsatellite markers. The YAC-based STS-content map of the region of chromosome 7 containing the human OB gene is depicted, as deduced by SEGMAP/Version 3.29 (Green and Green, 1991a; C.L. Magness and P. Green, unpublished data). The 19 uniquely-ordered STSs (see Table 1) are listed along the top. The 8 microsatellite-specific STSs are indicated with stars (see Table 2). Also indicated are the STSs corresponding to the PAX4 and OB genes as well as the predicted positions of the centromere (CEN) and 7q telomere (TEL) relative to the contig. Each of the 43 YAC clones is depicted by a horizontal bar, with its name given to the left and estimated YAC size (in kb, measured by pulsed-field gel electrophoresis) provided in parenthesis. The presence of an STS in a YAC is indicated by a darkened circle at the appropriate position. When an STS corresponds to the insert end of a YAC, a square is placed around the corresponding circle, both along the top (near the STS name) and at the end of the YAC from which it was derived. For the 5 YACs at the bottom (below the horizontal dashed line), 1 or more STS(s) expected to be present (based on the established STS order) was not detected [as assessed by testing the individual YACs with the corresponding STS-specific

PCR assay(s) at least twice], and these are depicted as open circles at the appropriate positions. Most of the YACs were isolated from a human-hamster hybrid cell-derived library (Green et al. 1995a), with their original names as indicated. The remaining YACs were isolated from total human genomic libraries, and their original library locations are provided in Table 3. Boxes are placed around the names of the 3 YACs (yWSS691, yWSS999, and yWSS2935) that were found by FISH analysis to map to 7q31.3. The contig is displayed in its uncomputed form, where YAC sizes are not used to estimate clone overlaps or STS spacing, and all of the STSs are therefore spaced in an equidistant fashion. In the computed form, where YAC sizes are used to estimate the relative distance separating each pair of adjacent STSs as well as the extent of clone overlaps, the total YAC contig appears to span just over 2 Mb.

THE REMAINDER OF THIS PAGE IS INTENTIONALLY LEFT BLANK

DETAILED DESCRIPTION

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual,* Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook et al., 1989"); *DNA Cloning: A Practical Approach,* Volumes I and II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* [B.D. Hames & S.J. Higgins eds. (1985)]; *Transcription And Translation* [B.D. Hames & S.J. Higgins, eds. (1984)]; *Animal Cell Culture* [R.I. Freshney, ed. (1986)]; *Immobilized Cells And Enzymes* [IRL Press, (1986)]; B. Perbal, *A Practical Guide To Molecular Cloning* (1984). Of particular relevance to the present invention are strategies for isolating, cloning, sequencing, analyzing, and characterizing a gene or nucleic acid based on the well known polymerase chain reaction (PCR) techniques.

Therefore, if appearing herein, the following terms shall have the definitions set out below.

The term "body weight modulator", "modulator", "modulators", and any variants not specifically listed, may be used herein interchangeably, and as used throughout the present application and claims refers in one instance to both nucleotides and to proteinaceous material, the latter including both single or multiple proteins. More specifically, the aforementioned terms extend to the nucleotides and to the DNA having the sequences described herein and presented in Figure 1 (SEQ ID NO:1), and Figure 2 (SEQ ID NO:3). Likewise, the proteins having the amino acid sequence data described herein and presented in Figure 1 (SEQ ID NO:2), and Figure 3 (SEQ ID NO:4) are likewise contemplated, as are the profile of activities set forth with respect to all materials both herein and in the claims. Accordingly, nucleotides displaying substantially equivalent or altered activity are likewise

contemplated, including substantially homologous analogs and allelic variations. Likewise, proteins displaying substantially equivalent or altered activity, including proteins modified deliberately, as for example, by site-directed mutagenesis, or accidentally through mutations in hosts that produce the modulators are likewise contemplated.

The terms "protein," which refers to the naturally occurring polypeptide, and "polypeptide" are used herein interchangeably with respect to the *ob* gene product and variants thereof. The term "mature protein" or "mature polypeptide" refers to the *ob* gene product with the signal sequence (or a fusion protein partner) removed.

A "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*, *i.e.*, capable of replication under its own control.

A "vector" is a replicon, such as a plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

A "cassette" refers to a segment of DNA that can be inserted into a vector at specific restriction sites. The segment of DNA encodes a polypeptide of interest, and the cassette and restriction sites are designed to ensure insertion of the cassette in the proper reading frame for transcription and translation.

"Heterologous" DNA refers to DNA not naturally located in the cell, or in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell.

A cell has been "transfected" by exogenous or heterologous DNA when such DNA has been introduced inside the cell. A cell has been "transformed" by exogenous

or heterologous DNA when the transfected DNA effects a phenotypic change. Preferably, the transforming DNA should be integrated (covalently linked) into chromosomal DNA making up the genome of the cell.

A "clone" is a population of cells derived from a single cell or common ancestor by mitosis.

A "nucleic acid molecule" refers to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules") in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary or quaternary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear or circular DNA molecules (*e.g.*, restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5′ to 3′ direction along the nontranscribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (*see* Sambrook et al., *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a T_m of 55°, can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x

SSC, 0.5% SDS). Moderate stringency hybridization conditions correspond to a higher T_m, *e.g.*, 40% formamide, with 5x or 6x SCC. High stringency hybridization conditions correspond to the highest T_m, *e.g.*, 50% formamide, 5x or 6x SCC. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T_m have been derived (*see* Sambrook et al., *supra*, 9.50-0.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (*see* Sambrook et al., *supra*, 11.7-11.8). Preferably a minimum length for a hybridizable nucleic acid is at least about 10 nucleotides; more preferably at least about 15 nucleotides; most preferably the length is at least about 20 nucleotides.

"Homologous recombination" refers to the insertion of a foreign DNA sequence of a vector in a chromosome. Preferably, the vector targets a specific chromosomal site for homologous recombination. For specific homologous recombination, the vector will contain sufficiently long regions of homology to sequences of the chromosome to allow complementary binding and incorporation of the vector into the chromosome. Longer regions of homology, and greater degrees of sequence similarity, may increase the efficiency of homologous recombination.

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in a cell *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of

the coding sequence are determined by a start codon at the 5′ (amino) terminus and a translation stop codon at the 3′ (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (*e.g.*, mammalian) DNA, and even synthetic DNA sequences. If the coding sequence is intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3′ to the coding sequence.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are control sequences.

A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced and translated into the protein encoded by the coding sequence.

A "signal sequence" is included at the beginning of the coding sequence of a protein to be expressed on the surface of a cell. This sequence encodes a signal peptide, N-terminal to the mature polypeptide, that directs the host cell to translocate the polypeptide. The term "translocation signal sequence" is also used herein to refer to this sort of signal sequence. Translocation signal sequences can be found associated with a variety of proteins native to eukaryotes and prokaryotes, and are often functional in both types of organisms.

A DNA sequence is "operatively linked" to an expression control sequence when the expression control sequence controls and regulates the transcription and translation of that DNA sequence. The term "operatively linked" includes having an appropriate start signal (*e.g.*, ATG) in front of the DNA sequence to be expressed and maintaining the correct reading frame to permit expression of the

DNA sequence under the control of the expression control sequence and production of the desired product encoded by the DNA sequence. If a gene that one desires to insert into a recombinant DNA molecule does not contain an appropriate start signal, such a start signal can be inserted upstream (5′) of and in reading frame with the gene.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3′ direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3′ terminus by the transcription initiation site and extends upstream (5′ direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

The term "standard hybridization conditions" refers to salt and temperature conditions substantially equivalent to 5 x SSC and 65°C for both hybridization and wash.

A molecule is "antigenic" when it is capable of specifically interacting with an antigen recognition molecule of the immune system, such as an immunoglobulin (antibody) or T cell antigen receptor. An antigenic polypeptide contains at least about 5, and preferably at least about 10, amino acids. An antigenic portion of a molecule can be that portion that is immunodominant for antibody or T cell receptor recognition, or it can be a portion used to generate an antibody to the molecule by conjugating the antigenic portion to a carrier molecule for immunization. A molecule that is antigenic need not be itself immunogenic, *i.e.*, capable of eliciting an immune response without a carrier.

An "antibody" is any immunoglobulin, including antibodies and fragments thereof, that binds a specific epitope. The term encompasses polyclonal, monoclonal, and chimeric antibodies, the last mentioned described in further detail in U.S. Patent Nos. 4,816,397 and 4,816,567, as well as antigen binding portions of antibodies, including Fab, $F(ab')_2$ and Fr (including single chain antibodies). Accordingly, the phrase "antibody molecule" in its various grammatical forms as used herein contemplates both an intact immunoglobulin molecule and an immunologically active portion of an immunoglobulin molecule containing the antibody combining site. An "antibody combining site" is that structural portion of an antibody molecule comprised of heavy and light chain variable and hypervariable regions that specifically binds antigen.

Exemplary antibody molecules are intact immunoglobulin molecules, substantially intact immunoglobulin molecules and those portions of an immunoglobulin molecule that contains the paratope, including those portions known in the art as Fab, Fab', $F(ab')_2$ and F(v), which portions are preferred for use in the therapeutic methods described herein.

Fab and $F(ab')_2$ portions of antibody molecules are prepared by the proteolytic reaction of papain and pepsin, respectively, on substantially intact antibody molecules by methods that are well-known. See for example, U.S. Patent No. 4,342,566 to Theofilopolous et al. Fab' antibody molecule portions are also well-known and are produced from $F(ab')_2$ portions followed by reduction of the disulfide bonds linking the two heavy chain portions as with mercaptoethanol, and followed by alkylation of the resulting protein mercaptan with a reagent such as iodoacetamide. An antibody containing intact antibody molecules is preferred herein.

The phrase "monoclonal antibody" in its various grammatical forms refers to an antibody having only one species of antibody combining site capable of immunoreacting with a particular antigen. A monoclonal antibody thus typically

displays a single binding affinity for any antigen with which it immunoreacts. A monoclonal antibody may therefore contain an antibody molecule having a plurality of antibody combining sites, each immunospecific for a different antigen; e.g., a bispecific (chimeric) monoclonal antibody.

A composition comprising "A" (where "A" is a single protein, DNA molecule, vector, recombinant host cell, etc.) is substantially free of "B" (where "B" comprises one or more contaminating proteins, DNA molecules, vectors, etc., but excluding racemic forms of A) when at least about 75% by weight of the proteins, DNA, vectors (depending on the category of species to which A and B belong) in the composition is "A". Preferably, "A" comprises at least about 90% by weight of the A+B species in the composition, most preferably at least about 99% by weight. It is also preferred that a composition, which is substantially free of contamination, contain only a single molecular weight species having the activity or characteristic of the species of interest.

The phrase "pharmaceutically acceptable" refers to molecular entities and compositions that are physiologically tolerable and do not typically produce an allergic or similar untoward reaction, such as gastric upset, dizziness and the like, when administered to a human. Preferably, as used herein, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the compound is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water or aqueous solution saline solutions and aqueous dextrose and glycerol solutions are preferably employed as carriers, particularly for injectable solutions. Suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin.

The phrase "therapeutically effective amount" is used herein to mean an amount sufficient to reduce by at least about 15 percent, preferably by at least 50 percent, more preferably by at least 90 percent, and most preferably prevent, a clinically significant deficit in the activity, function and response of the host. Alternatively, a therapeutically effective amount is sufficient to cause an improvement in a clinically significant condition in the host.

The term "adjuvant" refers to a compound or mixture that enhances the immune response to an antigen. An adjuvant can serve as a tissue depot that slowly releases the antigen and also as a lymphoid system activator that non-specifically enhances the immune response (Hood et al., *Immunology, Second Ed.*, 1984, Benjamin/Cummings: Menlo Park, California, p. 384). Often, a primary challenge with an antigen alone, in the absence of an adjuvant, will fail to elicit a humoral or cellular immune response. Adjuvants include, but are not limited to, complete Freund's adjuvant, incomplete Freund's adjuvant, saponin, mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil or hydrocarbon emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum*. Preferably, the adjuvant is pharmaceutically acceptable.

In its primary aspect, the present invention is directed to the identification of materials that function as modulators of mammalian body weight. In particular, the invention concerns the isolation, purification and sequencing of certain nucleic acids that correspond to the *ob* gene in both mice and humans, as well as the corresponding polypeptides expressed by these nucleic acids. The invention thus comprises the discovery of nucleic acids having the nucleotide sequences set forth in FIGURE 1 (SEQ ID NO:1) and FIGURE 2 (SEQ ID NO:3), and to degenerate variants, alleles and fragments thereof, all possessing the activity of modulating body weight and adiposity. The correspondence of the present nucleic acids to the *ob* gene portends their significant impact on conditions such as obesity as well as

other maladies and dysfunctions where abnormalities in body weight are a contributory factor. The invention extends to the proteins expressed by the nucleic acids of the invention, and particularly to those proteins set forth in FIGURE 1 (SEQ ID NO:2), FIGURE 3 (SEQ ID NO:4), FIGURE 5 (SEQ ID NO:5), and FIGURE 6 (SEQ ID NO:6), as well as conserved variants, active fragments, and cognate small molecules.

In particular, the present invention contemplates that naturally occurring fragments of the ob polypeptide may be important. The peptide sequence includes a number of sites that are frequently the target for proteolytic cleavage, *e.g.*, arginine residues. It is possible that the full length polypeptide may be cleaved at one or more such sites to form biologically active fragments. Such biologically active fragments may either agonize or antagonize the functional activity of the ob polypeptide to reduce body weight.

As discussed earlier, the weight control modulator peptides or their binding partners or other ligands or agents exhibiting either mimicry or antagonism to them or control over their production, may be prepared in pharmaceutical compositions, with a suitable carrier and at a strength effective for administration by various means to a patient experiencing abnormal fluctuations in body weight or adiposity, either alone or as part of an adverse medical condition such as cancer or AIDS, for the treatment thereof. A variety of administrative techniques may be utilized, among them parenteral techniques such as subcutaneous, intravenous and intraperitoneal injections, catheterizations and the like. Average quantities of the recognition factors or their subunits may vary and in particular should be based upon the recommendations and prescription of a qualified physician or veterinarian.

Also, antibodies including both polyclonal and monoclonal antibodies, and drugs that modulate the production or activity of the weight control modulators recognition factors and/or their subunits may possess certain diagnostic

applications and may for example, be utilized for the purpose of detecting and/or measuring conditions where abnormalities in body weight are or may be likely to develop. For example, the modulator peptides or their active fragments may be used to produce both polyclonal and monoclonal antibodies to themselves in a variety of cellular media, by known techniques such as the hybridoma technique utilizing, for example, fused mouse spleen lymphocytes and myeloma cells. These techniques are described in detail below. Likewise, small molecules that mimic or antagonize the activity(ies) of the receptor recognition factors of the invention may be discovered or synthesized, and may be used in diagnostic and/or therapeutic protocols.

Panels of monoclonal antibodies produced against modulator peptides can be screened for various properties; i.e., isotype, epitope, affinity, etc. Of particular interest are monoclonal antibodies that neutralize the activity of the modulator peptides. Such monoclonals can be readily identified in activity assays for the weight modulators. High affinity antibodies are also useful when immunoaffinity purification of native or recombinant modulator is possible.

Preferably, the anti-modulator antibody used in the diagnostic and therapeutic methods of this invention is an affinity purified polyclonal antibody. More preferably, the antibody is a monoclonal antibody (mAb). In addition, it is preferable for the anti-modulator antibody molecules used herein be in the form of Fab, Fab', $F(ab')_2$ or F(v) portions of whole antibody molecules.

As suggested earlier, a diagnostic method useful in the present invention comprises examining a cellular sample or medium by means of an assay including an effective amount of an antagonist to a modulator protein, such as an anti-modulator antibody, preferably an affinity-purified polyclonal antibody, and more preferably a mAb. In addition, it is preferable for the anti-modulator antibody molecules used herein be in the form of Fab, Fab', $F(ab')_2$ or F(v) portions or whole antibody molecules. As previously discussed, patients capable of benefiting

from this method include those suffering from cancer, AIDS, obesity or other condition where abnormal body weight is a characteristic or factor. Methods for isolating the modulator and inducing anti-modulator antibodies and for determining and optimizing the ability of anti-modulator antibodies to assist in the examination of the target cells are all well-known in the art.

The nucleic acids contemplated by the present invention extend as indicated, to other nucleic acids that code on expression for peptides such as those set forth in FIGURE 1 (SEQ ID NO:2), FIGURE 3 (SEQ ID NO:4), FIGURE 5 (SEQ ID NO:5), and FIGURE 6 (SEQ ID NO:6) herein. Accordingly, while specific DNA has been isolated and sequenced in relation to the *ob* gene, any animal cell potentially can serve as the nucleic acid source for the molecular cloning of a gene encoding the peptides of the invention. The DNA may be obtained by standard procedures known in the art from cloned DNA (*e.g.*, a DNA "library"), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (See, for example, Sambrook et al., 1989, *supra*; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II). Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will not contain intron sequences. Whatever the source, the gene should be molecularly cloned into a suitable vector for propagation of the gene.

In the molecular cloning of the gene from genomic DNA, the genomic DNA can be amplified using primers selected from the cDNA sequences. Alternatively, DNA fragments are generated, some of which will encode the desired gene. The DNA may be cleaved at specific sites using various restriction enzymes. One may also use DNAse in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The linear DNA fragments can then be separated according to size by standard techniques,

including but not limited to, agarose and polyacrylamide gel electrophoresis and column chromatography.

Once the DNA fragments are generated, identification of the specific DNA fragment containing the desired *ob* or *ob*-like gene may be accomplished in a number of ways. For example, if an amount of a portion of a *ob* or *ob*-like gene or its specific RNA, or a fragment thereof, is available and can be purified and labeled, the generated DNA fragments may be screened by nucleic acid hybridization to the labeled probe (Benton and Davis, 1977, *Science* **196**:180; Grunstein and Hogness, 1975, *Proc. Natl. Acad. Sci. U.S.A.* **72**:3961). The present invention provides such nucleic acid probes, which can be conveniently prepared from the specific sequences disclosed herein, *e.g.*, a hybridizable probe having a nucleotide sequence corresponding to at least a 10, and preferably a 15, nucleotide fragment of the sequences depicted in Figure 1 (SEQ ID NO:1) or Figure 2 (SEQ ID NO:3). Preferably, a fragment is selected that is highly unique to the modulator peptides of the invention. Those DNA fragments with substantial homology to the probe will hybridize. As noted above, the greater the degree of homology, the more stringent hybridization conditions can be used. In one embodiment, low stringency hybridization conditions are used to identify a homologous modulator peptide. However, in a preferred aspect, and as demonstrated experimentally herein, a nucleic acid encoding a modulator peptide of the invention will hybridize to a nucleic acid having a nucleotide sequence such as depicted in Figure 1 (SEQ ID NO:1) or Figure 2 (SEQ ID NO:3), or a hybridizable fragment thereof, under moderately stringent conditions; more preferably, it will hybridize under high stringency conditions.

Alternatively, the presence of the gene may be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones, or DNA clones which hybrid-select the proper mRNAs, can be selected which produce a protein that, *e.g.*, has similar or identical electrophoretic migration, isoelectric focusing behavior, proteolytic digestion

maps, tyrosine phosphatase activity or antigenic properties as known for the present modulator peptides. For example, the antibodies of the instant invention can conveniently be used to screen for homologs of modulator peptides from other sources.

A gene encoding a modulator peptide of the invention can also be identified by mRNA selection, *i.e.*, by nucleic acid hybridization followed by *in vitro* translation. In this procedure, fragments are used to isolate complementary mRNAs by hybridization. Such DNA fragments may represent available, purified modulator DNA. Immunoprecipitation analysis or functional assays (*e.g.*, tyrosine phosphatase activity) of the *in vitro* translation products of the products of the isolated mRNAs identifies the mRNA and, therefore, the complementary DNA fragments, that contain the desired sequences. In addition, specific mRNAs may be selected by adsorption of polysomes isolated from cells to immobilized antibodies specifically directed against a modulator peptide.

A radiolabeled modulator peptide cDNA can be synthesized using the selected mRNA (from the adsorbed polysomes) as a template. The radiolabeled mRNA or cDNA may then be used as a probe to identify homologous modulator peptide DNA fragments from among other genomic DNA fragments.

Another feature of this invention is the expression of the DNA sequences disclosed herein. As is well known in the art, DNA sequences may be expressed by operatively linking them to an expression control sequence in an appropriate expression vector and employing that expression vector to transform an appropriate unicellular host.

Such operative linking of a DNA sequence of this invention to an expression control sequence, of course, includes, if not already part of the DNA sequence, the provision of an initiation codon, ATG, in the correct reading frame upstream of the DNA sequence.

A wide variety of host/expression vector combinations may be employed in expressing the DNA sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and Synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, e.g., *E. coli* plasmids col El, pCR1, pBR322, pMB9, pUC or pUC plasmid derivatives, *e.g.*, pGEX vectors, pET vectors, pmal-c, pFLAG, etc., and their derivatives, plasmids such as RP4; phage DNAs, e.g., the numerous derivatives of phage λ, e.g., NM989, and other phage DNA, e.g., M13 and Filamentous single stranded phage DNA; yeast plasmids such as the 2μ plasmid or derivatives thereof; vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences; and the like. In a preferred embodiment, expression of *ob* is achieved in methylotrophic yeast, *e.g.*, *Pichia pastoris* yeast (*see, e.g.*, International Patent Publication No. WO 90/03431, published 5 April 1990, by Brierley et al.; International Patent Publication No. WO 90/10697, published 20 September 1990, by Siegel et al.). In a specific embodiment, *infra*, an expression vector is engineered for expression of *ob* under control of the α-mating factor signal sequence.

Any of a wide variety of expression control sequences -- sequences that control the expression of a DNA sequence operatively linked to it -- may be used in these vectors to express the DNA sequences of this invention. Such useful expression control sequences include, for example, the early or late promoters of SV40, CMV, vaccinia, polyoma or adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the *LTR* system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase (e.g., Pho5), the AOX 1 promoter of methylotrophic yeast, the promoters of the yeast α-mating factors, and other sequences known to control the

expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

A wide variety of unicellular host cells are also useful in expressing the DNA sequences of this invention. These hosts may include well known eukaryotic and prokaryotic hosts, such as strains of *E. coli*, *Pseudomonas*, *Bacillus*, *Streptomyces*; fungi such as yeasts (*Saccharomyces*, and methylotrophic yeast such as *Pichia*, *Candida*, *Hansenula*, and *Torulopsis*); and animal cells, such as CHO, Rl.l, B-W and L-M cells, African Green Monkey kidney cells (e.g., COS 1, COS 7, BSC1, BSC40, and BMT10), insect cells (e.g., Sf9), and human cells and plant cells in tissue culture.

It will be understood that not all vectors, expression control sequences and hosts will function equally well to express the DNA sequences of this invention. Neither will all hosts function equally well with the same expression system. However, one skilled in the art will be able to select the proper vectors, expression control sequences, and hosts without undue experimentation to accomplish the desired expression without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must function in it. The vector's copy number, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, will also be considered.

In selecting an expression control sequence, a variety of factors will normally be considered. These include, for example, the relative strength of the system, its controllability, and its compatibility with the particular DNA sequence or gene to be expressed, particularly as regards potential secondary structures. Suitable unicellular hosts will be selected by consideration of, e.g., their compatibility with the chosen vector, their secretion characteristics, their ability to fold proteins correctly, and their fermentation requirements, as well as the toxicity to the host

of the product encoded by the DNA sequences to be expressed, and the ease of purification of the expression products.

Considering these and other factors a person skilled in the art will be able to construct a variety of vector/expression control sequence/host combinations that will express the DNA sequences of this invention on fermentation or in large scale animal culture.

In a specific embodiment, an ob fusion protein can be expressed. An ob fusion protein comprises at least a functionally active portion of a non-ob protein joined via a peptide bond to at least a functionally active portion of an ob polypeptide. The non-ob sequences can be amino- or carboxy-terminal to the ob sequences. More preferably, for stable expression of a proteolytically inactive ob fusion protein, the portion of the non-ob fusion protein is joined via a peptide bond to the amino terminus of the ob protein. A recombinant DNA molecule encoding such a fusion protein comprises a sequence encoding at least a functionally active portion of a non-ob protein joined in-frame to the ob coding sequence, and preferably encodes a cleavage site for a specific protease, *e.g.*, thrombin or Factor Xa, preferably at the ob-non-ob juncture. In a specific embodiment, the fusion protein is expressed in *Escherichia coli* or in *P. pastoris*.

In a specific embodiment, *infra*, vectors were prepared to express the murine and human *ob* genes, with and without the codon for gln-49, in bacterial expression systems and yeast (*Pichia*) expression systems as fusion proteins. The *ob* gene is prepared with an endonuclease cleavage site, *e.g.*, using PCR and novel primers. It is desirable to confirm sequences generated by PCR, since the probability of including a point mutation is greater with this technique. A plasmid containing a histidine tag (HIS-TAG) and a proteolytic cleavage site is used. The presence of the histidine makes possible the selective isolation of recombinant proteins on a Ni-chelation column, or by affinity purification. The proteolytic cleavage site, in a specific embodiment, *infra*, a thrombin cleavage site, is engineered so that

treatment with the protease, *e.g.*, thrombin, will release the full length mature (*i.e.*, lacking a signal sequence) ob polypeptide.

In another aspect, the pGEX vector (Smith and Johnson, 1988, Gene 67:31-40) can be used. This vector fuses the schistosoma japonicum glutathionine S-transferase cDNA to the sequence of interest. Bacterial proteins are harvested and recombinant proteins can be quickly purified on a reduced glutathione affinity column. The GST carrier can subsequently be cleaved from fusion proteins by digestion with site-specific proteases. After cleavage, the carrier and uncleaved fusion protein can be removed by absorption on glutathione agarose. Difficulty with the system occasionally arises when the encoded protein is insoluble in aqueous solutions.

Expression of recombinant proteins in bacterial systems may result in incorrect folding of the expressed protein, requiring refolding. The recombinant protein can be refolded prior to or after cleavage to form a functionally active ob polypeptide. The ob polypeptide may be refolded by the steps of (i) incubating the protein in a denaturing buffer that contains a reducing agent, and then (ii) incubating the protein in a buffer that contains an oxidizing agent, and preferably also contains a protein stabilizing agent or a chaotropic agent, or both. Suitable redox (reducing/oxidizing) agent pairs include, but are not limited to, reduced glutathione/glutathione disulfide, cystine/cysteine, cystamine/cysteamine, and 2-mercaptoethanol/2-hydroxyethyldisulfide. In a particular aspect, the fusion protein can be solubilized in a denaturant, such as urea, prior to exchange into the reducing buffer. In preferred embodiment, the protein is also purified, *e.g.*, by ion exchange or Ni-chelation chromatography, prior to exchange into the reducing buffer. Denaturing agents include but are not limited to urea and guanidine-HCl. The recombinant protein is then diluted about at least 10-fold, more preferably about 100-fold, into an oxidizing buffer that contains an oxidizing agent, such as but not limited to 0.1 M Tris-HCl, pH 8.0, 1 mM EDTA, 0.15 M NaCl, 0.3 M oxidized glutathione. The fusion protein is then incubated for about 1 to about 24

hours, preferably about 2 to about 16 hours, at room temperature in the oxidizing buffer. The oxidizing buffer may comprise a protein stabilizing agent, *e.g.*, a sugar, an alcohol, or ammonium sulfate. The oxidizing buffer may further comprises a chaotropic agent at low concentration, to destabilize incorrect intermolecular interactions and thus promote proper folding. Suitable chaotropic agents include but are not limited to a detergent, a polyol, L-arginine, guanidine-HCl and polyethylene glycol (PEG). It is important to use a low enough concentration of the chaotropic agent to avoid denaturing the protein. The refolded protein can be concentrated by at least about 10-fold, more preferably by the amount it was diluted into the oxidizing buffer.

Bacterial fermentation processes can also result in a recombinant protein preparation that contains unacceptable levels of endotoxins. Therefore, the invention contemplates removal of such endotoxins, *e.g.*, by using endotoxin-specific antibodies or other endotoxin binding molecules. The presence of endotoxins can be determined by standard techniques, such as by employing E-TOXATE Reagents (Sigma), or with bioassays.

In addition to the specific example, the present inventors contemplate use of baculovirus, mammalian, and yeast expression systems to express the ob protein. For example, in baculovirus expression systems, both non-fusion transfer vectors, such as but not limited to pVL941 (BamH1 cloning site; Summers), pVL1393 (BamH1, SmaI, XbaI, EcoR1, NotI, XmaIII, BglII, and PstI cloning site; Invitrogen), pVL1392 (BglII, PstI, NotI, XmaIII, EcoRI, XbaI, SmaI, and BamH1 cloning site; Summers and Invitrogen), and pBlueBacIII (BamH1, BglII, PstI, NcoI, and HindIII cloning site, with blue/white recombinant screening possible; Invitrogen), and fusion transfer vectors, such as but not limited to pAc700 (BamH1 and KpnI cloning site, in which the BamH1 recognition site begins with the initiation codon; Summers), pAc701 and pAc702 (same as pAc700, with different reading frames), pAc360 (BamH1 cloning site 36 base pairs downstream of a polyhedrin initiation codon; Invitrogen(195)), and pBlueBacHisA, B, C (three

different reading frames, with BamH1, BglII, PstI, NcoI, and HindIII cloning site, an N-terminal peptide for ProBond purification, and blue/white recombinant screening of plaques; Invitrogen (220)).

Mammalian expression vectors contemplated for use in the invention include vectors with inducible promoters, such as dihydrofolate reductase (DHFR), *e.g.*, any expression vector with a DHFR expression vector, or a DHFR/methotrexate co-amplification vector, such as pED (PstI, SalI, SbaI, SmaI, and EcoRI cloning site, with the vector expressing both the cloned gene and DHFR; *see* Kaufman, *Current Protocols in Molecular Biology*, 16.12, 1991). Alternatively, a glutamine synthetase/methionine sulfoximine co-amplification vector, such as pEE14 (HindIII, XbaI, SmaI, SmaI, EcoRI, and BclI cloning site, in which the vector expresses glutamine synthase and the cloned gene; Celltech). In another embodiment, a vector that directs episomal expression under control of Epstein Barr Virus (EBV) can be used, such as pREP4 (BamH1, SfiI, XhoI, NotI, NheI, HindIII, NheI, PvuII, and KpnI cloning site, constitutive RSV LTR promoter, hygromycin selectable marker; Invitrogen), pCEP4 (BamH1, SfiI, XhoI, NotI, NheI, HindIII, NheI, PvuII, and KpnI cloning site, constitutive hCMV immediate early gene, hygromycin selectable marker; Invitrogen), pMEP4 (KpnI, PvuI, NheI, HindIII, NotI, XhoI, SfiI, BamH1 cloning site, inducible methallothionein IIa gene promoter, hygromycin selectable marker: Invitrogen), pREP8 (BamH1, XhoI, NotI, HindIII, NheI, and KpnI cloning site, RSV LTR promoter, histidinol selectable marker; Invitrogen), pREP9 (KpnI, NheI, HindIII, NotI, XhoI, SfiI, and BamHI cloning site, RSV LTR promoter, G418 selectable marker; Invitrogen), and pEBVHis (RSV LTR promoter, hygromycin selectable marker, N-terminal peptide purifiable via ProBond resin and cleaved by enterokinase; Invitrogen). Selectable mammalian expression vectors for use in the invention include pRc/CMV (HindIII, BstXI, NotI, SbaI, and ApaI cloning site, G418 selection; Invitrogen), pRc/RSV (HindIII, SpeI, BstXI, NotI, XbaI cloning site, G418 selection; Invitrogen), and others. Vaccinia virus mammalian expression vectors (*see*, Kaufman, *supra*) for use according to the invention include but are not

limited to pSC11 (SmaI cloning site, TK- and β-gal selection), pMJ601 (SalI, SmaI, AflI, NarI, BspMII, BamHI, ApaI, NheI, SacII, KpnI, and HindIII cloning site; TK- and β-gal selection), and pTKgptF1S (EcoRI, PstI, SalI, AccI, HindII, SbaI, BamHI, and HpA cloning site, TK or XPRT selection).

Yeast expression systems can also be used according to the invention to express ob polypeptide. For example, the non-fusion pYES2 vector (XbaI, SphI, ShoI, NotI, GstXI, EcoRI, BstXI, BamH1, SacI, Kpn1, and HindIII cloning sit; Invitrogen) or the fusion pYESHisA, B, C (XbaI, SphI, ShoI, NotI, BstXI, EcoRI, BamH1, SacI, KpnI, and HindIII cloning site, N-terminal peptide purified with ProBond resin and cleaved with enterokinase; Invitrogen), to mention just two, can be employed according to the invention.

It is further intended that body weight modulator peptide analogs may be prepared from nucleotide sequences derived within the scope of the present invention. Analogs, such as fragments, may be produced, for example, by pepsin digestion of weight modulator peptide material. Other analogs, such as muteins, can be produced by standard site-directed mutagenesis of weight modulator peptide coding sequences. Analogs exhibiting "weight modulator activity" such as small molecules, whether functioning as promoters or inhibitors, may be identified by known *in vivo* and/or *in vitro* assays.

In addition to recombinant expression of ob polypeptide, the present invention envisions and fully enables preparation of ob polypeptide, or fragments thereof, using the well known and highly developed techniques of solid phase peptide synthesis. The invention contemplates using both the popular Boc and Fmoc, as well as other protecting group strategies, for preparing ob polypeptide or fragments thereof. Various techniques for refolding and oxidizing the cysteine side chains to form a disulfide bond are also well known in the art.

Derivatives of Ob Peptides

Generally, the present protein (herein the term "protein" is used to include "peptide", unless otherwise indicated) may be derivatized by the attachment of one or more chemical moieties to the protein moiety. The chemically modified derivatives may be further formulated for intraarterial, intraperitoneal, intramuscular subcutaneous, intravenous, oral, nasal, pulmonary, topical or other routes of administration. Chemical modification of biologically active proteins has been found to provide additional advantages under certain circumstances, such as increasing the stability and circulation time of the therapeutic protein and decreasing immunogenicity. See U.S. Patent No. 4,179,337, Davis et al., issued December 18, 1979. For a review, see Abuchowski et al., in Enzymes as Drugs. (J.S. Holcerberg and J. Roberts, eds. pp. 367-383 (1981)). A review article describing protein modification and fusion proteins is Francis, *Focus on Growth Factors* 3: 4-10 (May 1992) (published by Mediscript, Mountview Court, Friern Barnet Lane, London N20, OLD, UK).

Chemical Moieties For Derivatization

The chemical moieties suitable for derivatization may be selected from among water soluble polymers. The polymer selected should be water soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable. One skilled in the art will be able to select the desired polymer based on such considerations as whether the polymer/protein conjugate will be used therapeutically, and if so, the desired dosage, circulation time, resistance to proteolysis, and other considerations. For the present proteins and peptides, these may be ascertained using the assays provided herein.

Polymer Molecules

The water soluble polymer may be selected from the group consisting of, for example, polyethylene glycol, copolymers of ethylene glycol/propylene glycol,

carboxymethylcellulose, dextran, polyvinyl alcohol, polyvinyl pyrrolidone, poly-1, 3-dioxolane, poly-1,3,6-trioxane, ethylene/maleic anhydride copolymer, polyaminoacids (either homopolymers or random copolymers), and dextran or poly(n-vinyl pyrrolidone)polyethylene glycol, propropylene glycol homopolymers, prolypropylene oxide/ethylene oxide co- polymers, polyoxyethylated polyols and polyvinyl alcohol. Polyethylene glycol propionaldenhyde may advantages in manufacturing due to its stability in water.

The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 2kDa and about 100kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

Polymer/Protein Ratio

The number of polymer molecules so attached may vary, and one skilled in the art will be able to ascertain the effect on function. One may mono-derivatize, or may provide for a di-, tri-, tetra- or some combination of derivatization, with the same or different chemical moieties (e.g., polymers, such as different weights of polyethylene glycols). The proportion of polymer molecules to protein (or peptide) molecules will vary, as will their concentrations in the reaction mixture. In general, the optimum ratio (in terms of efficiency of reaction in that there is no excess unreacted protein or polymer) will be determined by factors such as the desired degree of derivatization (e.g., mono, di-, tri-, etc.), the molecular weight of the polymer selected, whether the polymer is branched or unbranched, and the reaction conditions.

Attachment of the Chemical Moiety to the Protein

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art. E.g., EP 0 401 384 herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., Exp. Hematol. 20: 1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N- terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydrl groups may also be used as a reactive group for attaching the polyethylene glycol molecule(s). Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group. Attachment at residues important for receptor binding should be avoided if receptor binding is desired.

N-terminally chemically modified proteins.

One may specifically desire N-terminally chemically modified protein. Using polyethylene glycol as an illustration of the present compositions, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (or peptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective N-terminal chemically modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in

a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved. For example, one may selectively N-terminally pegylate the protein by performing the reaction at pH which allows one to take advantage of the pK_a differences between the ϵ-amino groups of the lysine residues and that of the α-amino group of the N-terminal residue of the protein. By such selective derivatization attachment of a water soluble polymer to a protein is controlled: the conjugation with the polymer takes place predominantly at the N-terminus of the protein and no significant modification of other reactive groups, such as the lysine side chain amino groups, occurs. Using reductive alkylation, the water soluble polymer may be of the type described above, and should have a single reactive aldehyde for coupling to the protein. Polyethylene glycol proprionaldehyde, containing a single reactive aldehyde, may be used.

As mentioned above, a DNA sequence encoding weight modulator peptides as disclosed herein can be prepared synthetically rather than cloned. The DNA sequence can be designed with the appropriate codons for the weight modulator peptide amino acid sequences. In general, one will select preferred codons for the intended host if the sequence will be used for expression. The complete sequence is assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge, *Nature*, **292**:756 (1981); Nambair et al., *Science*, **223**:1299 (1984); Jay et al., *J. Biol. Chem.*, *259*:6311 (1984).

Synthetic DNA sequences allow convenient construction of genes which will express weight modulator analogs or "muteins". Alternatively, DNA encoding muteins can be made by site-directed mutagenesis of native modulator genes or cDNAs, and muteins can be made directly using conventional polypeptide synthesis.

A general method for site-specific incorporation of unnatural amino acids into proteins is described in Christopher J. Noren, Spencer J. Anthony-Cahill, Michael C. Griffith, Peter G. Schultz, *Science*, **244**:182-188 (April 1989). This method may be used to create analogs of the ob polypeptide with unnatural amino acids.

The present invention extends to the preparation of antisense nucleotides and ribozymes that may be used to interfere with the expression of the weight modulator proteins at the translational level. This approach utilizes antisense nucleic acid and ribozymes to block translation of a specific mRNA, either by masking that mRNA with an antisense nucleic acid or cleaving it with a ribozyme.

Antisense nucleic acids are DNA or RNA molecules that are complementary to at least a portion of a specific mRNA molecule (See Weintraub, 1990; Marcus-Sekura, 1988). In the cell, they hybridize to that mRNA, forming a double stranded molecule. The cell does not translate an mRNA in this double-stranded form. Therefore, antisense nucleic acids interfere with the expression of mRNA into protein. Oligomers of about fifteen nucleotides and molecules that hybridize to the AUG initiation codon will be particularly efficient, since they are easy to synthesize and are likely to pose fewer problems than larger molecules when introducing them into weight modulator peptide-producing cells. Antisense methods have been used to inhibit the expression of many genes *in vitro* (Marcus-Sekura, 1988; Hambor et al., 1988).

Ribozymes are RNA molecules possessing the ability to specifically cleave other single stranded RNA molecules in a manner somewhat analogous to DNA restriction endonucleases. Ribozymes were discovered from the observation that certain mRNAs have the ability to excise their own introns. By modifying the nucleotide sequence of these RNAs, researchers have been able to engineer molecules that recognize specific nucleotide sequences in an RNA molecule and cleave it (Cech, 1988). Because they are sequence-specific, only mRNAs with particular sequences are inactivated.

Investigators have identified two types of ribozymes, *Tetrahymena*-type and "hammerhead"-type (Hasselhoff and Gerlach, 1988). *Tetrahymena*-type ribozymes recognize four-base sequences, while "hammerhead"-type recognize eleven- to eighteen-base sequences. The longer the recognition sequence, the more likely it is to occur exclusively in the target MRNA species. Therefore, hammerhead-type ribozymes are preferable to *Tetrahymena*-type ribozymes for inactivating a specific mRNA species, and eighteen base recognition sequences are preferable to shorter recognition sequences.

The DNA sequences described herein may thus be used to prepare antisense molecules against and ribozymes that cleave mRNAs for weight modulator proteins and their ligands, thus inhibiting expression of the *ob* gene, and leading to increased weight gain and adiposity.

The present invention also relates to a variety of diagnostic applications, including methods for detecting the presence of conditions and/or stimuli that impact abnormalities in body weight or adiposity, by reference to their ability to elicit the activities which are mediated by the present weight modulators. As mentioned earlier, the weight modulator peptides can be used to produce antibodies to themselves by a variety of known techniques, and such antibodies could then be isolated and utilized as in tests for the presence of particular transcriptional activity in suspect target cells.

Antibody(ies) to the body weight modulators, *i.e.*, the ob polypeptide, can be produced and isolated by standard methods including the well known hybridoma techniques. For convenience, the antibody(ies) to the weight modulators will be referred to herein as Ab_1 and antibody(ies) raised in another species as Ab_2.

According to the invention, ob polypeptide produced recombinantly or by chemical synthesis, and fragments or other derivatives or analogs thereof, including fusion proteins, may be used as an immunogen to generate antibodies that recognize the

ob polypeptide. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library.

Various procedures known in the art may be used for the production of polyclonal antibodies to ob polypeptide a recombinant PTP or derivative or analog thereof. For the production of antibody, various host animals can be immunized by injection with the ob polypeptide, or a derivative (*e.g.*, fragment or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the ob polypeptide or fragment thereof can be conjugated to an immunogenic carrier, *e.g.*, bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum.*

For preparation of monoclonal antibodies directed toward the ob polypeptide, or fragment, analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilizing recent technology (PCT/US90/02545). According to the invention, human antibodies may be used and can be obtained by using human hybridomas (Cote et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030) or by transforming human

B cells with EBV virus *in vitro* (Cole et al., 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, pp. 77-96). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, J. Bacteriol. 159-870; Neuberger et al., 1984, Nature 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule specific for an ob polypeptide together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention. Such human or humanized chimeric antibodies are preferred for use in therapy of human diseases or disorders (described *infra*), since the human or humanized antibodies are much less likely than xenogenic antibodies to induce an immune response, in particular an allergic response, themselves.

According to the invention, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce ob polypeptide-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse et al., 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for an ob polypeptide, or its derivatives, or analogs.

Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.*, radioimmunoassay, ELISA (enzyme-linked immunosorbent assay), "sandwich" immunoassays,

immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of an ob polypeptide, one may assay generated hybridomas for a product which binds to an ob polypeptide fragment containing such epitope. For selection of an antibody specific to an ob polypeptide from a particular species of animal, one can select on the basis of positive binding with ob polypeptide expressed by or isolated from cells of that species of animal.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the ob polypeptide, *e.g.*, for Western blotting, imaging ob polypeptide *in situ*, measuring levels thereof in appropriate physiological samples, etc.

In a specific embodiment, antibodies that agonize or antagonize the activity of ob polypeptide can be generated. Such antibodies can be tested using the assays described *infra* for identifying ligands.

Immortal, antibody-producing cell lines can also be created by techniques other than fusion, such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, e.g., M. Schreier et al., "Hybridoma Techniques" (1980); Hammerling et al., "Monoclonal Antibodies And T-cell Hybridomas" (1981); Kennett et al., "Monoclonal Antibodies" (1980); see also

U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,451,570; 4,466,917; 4,472,500; 4,491,632; 4,493,890.

In a specific embodiment, antibodies are developed by immunizing rabbits with synthetic peptides predicted by the protein sequence or with recombinant proteins made using bacterial expression vectors. The choice of synthetic peptides is made after careful analysis of the predicted protein structure, as described above. In particular, peptide sequences between putative cleavage sites are chosen. Synthetic peptides are conjugated to a carrier such as KLH hemocyanin or BSA using carbodiimide and used in Freunds adjuvant to immunize rabbits. In order to prepare recombinant protein, the *gex* vector can be used to express the polypeptide (Smith and Johnson, *supra*). Alternatively, one can use only hydrophilic domains to generate the fusion protein. The expressed protein will be prepared in quantity and used to immunize rabbits in Freunds adjuvant.

The presence of weight modulator in cells can be ascertained by the usual immunological procedures applicable to such determinations. A number of useful procedures are known. Three such procedures which are especially useful utilize either the receptor recognition factor labeled with a detectable label, antibody Ab_1 labeled with a detectable label, or antibody Ab_2 labeled with a detectable label. The procedures may be summarized by the following equations wherein the asterisk indicates that the particle is labeled, and "WM" stands for the weight modulator:

A. $WM^* + Ab_1 = WM^*Ab_1$

B. $WM + Ab^* = WMAb_1^*$

C. $WM + Ab_1 + Ab_2^* = Ab_1WMAb_2^*$

The procedures and their application are all familiar to those skilled in the art and accordingly may be utilized within the scope of the present invention. The "competitive" procedure, Procedure A, is described in U.S. Patent Nos. 3,654,090 and 3,850,752. Procedure B is representative of the well known competitive assay

techniques. Procedure C, the "sandwich" procedure, is described in U.S. Patent Nos. RE 31,006 and 4,016,043. Still other procedures are known such as the "double antibody", or "DASP" procedure.

In each instance, the weight modulators form complexes with one or more antibody(ies) or binding partners and one member of the complex is labeled with a detectable label. The fact that a complex has formed and, if desired, the amount thereof, can be determined by known methods applicable to the detection of labels.

It will be seen from the above, that a characteristic property of Ab_2 is that it will react with Ab_1. This is because Ab_1 raised in one mammalian species has been used in another species as an antigen to raise the antibody Ab_2. For example, Ab_2 may be raised in goats using rabbit antibodies as antigens. Ab_2 therefore would be anti-rabbit antibody raised in goats. For purposes of this description and claims, Ab_1 will be referred to as a primary or anti-weight modulator antibody, and Ab_2 will be referred to as a secondary or anti-Ab_1 antibody.

The labels most commonly employed for these studies are radioactive elements, enzymes, chemicals which fluoresce when exposed to ultraviolet light, and others.

A number of fluorescent materials are known and can be utilized as labels. These include, for example, fluorescein, rhodamine and auramine. A particular detecting material is anti-rabbit antibody prepared in goats and conjugated with fluorescein through an isothiocyanate.

The weight modulators or their binding partners can also be labeled with a radioactive element or with an enzyme. The radioactive label can be detected by any of the currently available counting procedures. The preferred isotope may be selected from ^{3}H, ^{14}C, ^{32}P, ^{35}S, ^{36}Cl, ^{51}Cr, ^{57}Co, ^{58}Co, ^{59}Fe, ^{90}Y, ^{125}I, ^{131}I, and ^{186}Re.

Enzyme labels are likewise useful, and can be detected by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques. The enzyme is conjugated to the selected particle by reaction with bridging molecules such as carbodiimides, diisocyanates, glutaraldehyde and the like. Many enzymes which can be used in these procedures are known and can be utilized. The preferred are peroxidase, ß-glucuronidase, ß-D-glucosidase, ß-D-galactosidase, urease, glucose oxidase plus peroxidase and alkaline phosphatase. U.S. Patent Nos. 3,654,090; 3,850,752; and 4,016,043 are referred to by way of example for their disclosure of alternate labeling material and methods.

A particular assay system that is to be utilized in accordance with the present invention, is known as a receptor assay. In a receptor assay, the material to be assayed is appropriately labeled and then certain cellular test colonies are inoculated with a quantity of both the labeled and unlabeled material after which binding studies are conducted to determine the extent to which the labeled material binds to the cell receptors. In this way, differences in affinity between materials can be ascertained.

Accordingly, a purified quantity of the weight modulator may be radiolabeled and combined, for example, with antibodies or other inhibitors thereto, after which binding studies would be carried out. Solutions would then be prepared that contain various quantities of labeled and unlabeled uncombined weight modulator, and cell samples would then be inoculated and thereafter incubated. The resulting cell monolayers are then washed, solubilized and then counted in a gamma counter for a length of time sufficient to yield a standard error of $<5\%$. These data are then subjected to Scatchard analysis after which observations and conclusions regarding material activity can be drawn. While the foregoing is exemplary, it illustrates the manner in which a receptor assay may be performed and utilized, in the instance where the cellular binding ability of the assayed material may serve as a distinguishing characteristic. In turn, a receptor assay will be particularly useful

in the identification of the specific receptors to the present modulators, such as the *db* receptor.

A further assay useful and contemplated in accordance with the present invention is known as a "cis/trans" assay. Briefly, this assay employs two genetic constructs, one of which is typically a plasmid that continually expresses a particular receptor of interest when transfected into an appropriate cell line, and the second of which is a plasmid that expresses a reporter such as luciferase, under the control of a receptor/ligand complex. Thus, for example, if it is desired to evaluate a compound as a ligand for a particular receptor, one of the plasmids would be a construct that results in expression of the receptor in the chosen cell line, while the second plasmid would possess a promoter linked to the luciferase gene in which the response element to the particular receptor is inserted. If the compound under test is an agonist for the receptor, the ligand will complex with the receptor, and the resulting complex will bind the response element and initiate transcription of the luciferase gene. The resulting chemiluminescence is then measured photometrically, and dose response curves are obtained and compared to those of known ligands. The foregoing protocol is described in detail in U.S. Patent No. 4,981,784 and PCT International Publication No. WO 88/03168, for which purpose the artisan is referred.

In a further embodiment of this invention, commercial test kits suitable for use by a medical specialist may be prepared to determine the presence or absence of predetermined transcriptional activity or predetermined transcriptional activity capability in suspected target cells. In accordance with the testing techniques discussed above, one class of such kits will contain at least the labeled weight modulator or its binding partner, for instance an antibody specific thereto, and directions, of course, depending upon the method selected, e.g., "competitive", "sandwich", "DASP" and the like. The kits may also contain peripheral reagents such as buffers, stabilizers, etc.

Accordingly, a test kit may be prepared for the demonstration of the presence or capability of cells for predetermined transcriptional activity, comprising:

(a) a predetermined amount of at least one labeled immunochemically reactive component obtained by the direct or indirect attachment of the present weight modulator or a specific binding partner thereto, to a detectable label;

(b) other reagents; and

(c) directions for use of said kit.

More specifically, the diagnostic test kit may comprise:

(a) a known amount of the weight modulator as described above (or a binding partner) generally bound to a solid phase to form an immunosorbent, or in the alternative, bound to a suitable tag, or plural such end products, etc. (or their binding partners) one of each;

(b) if necessary, other reagents; and

(c) directions for use of said test kit.

In a further variation, the test kit may be prepared and used for the purposes stated above, which operates according to a predetermined protocol (e.g. "competitive", "sandwich", "double antibody", etc.), and comprises:

(a) a labeled component which has been obtained by coupling the weight modulator to a detectable label;

(b) one or more additional immunochemical reagents of which at least one reagent is a ligand or an immobilized ligand, which ligand is selected from the group consisting of:

(i) a ligand capable of binding with the labeled component (a);

(ii) a ligand capable of binding with a binding partner of the labeled component (a);

(iii) a ligand capable of binding with at least one of the component(s) to be determined; and

(iv) a ligand capable of binding with at least one of the binding partners of at least one of the component(s) to be determined; and

(c) directions for the performance of a protocol for the detection and/or determination of one or more components of an immunochemical reaction between the weight modulator and a specific binding partner thereto.

In accordance with the above, an assay system for screening potential drugs effective to mimic or antagonize the activity of the weight modulator may be prepared. The weight modulator may be introduced into a test system, and the prospective drug may also be introduced into the resulting cell culture, and the culture thereafter examined to observe any changes in the activity of the cells, due either to the addition of the prospective drug alone, or due to the effect of added quantities of the known weight modulator.

As stated earlier, the molecular cloning of the *ob* gene described herein has led to the identification of a class of materials that function on the molecular level to modulate mammalian body weight. The discovery of the modulators of the invention has important implications for the diagnosis and treatment of nutritional disorders including, but not limited to, obesity, weight loss associated with cancer and the treatment of diseases associated with obesity such as hypertension, heart disease and Type II diabetes. In addition, there are potential agricultural uses for the gene product in cases where one might wish to modulate the body weight of domestic animals. Finally, to the extent that one or more of the modulators of the invention are secreted molecules, they can be used biochemically to isolate their receptor using the technology of expression cloning. The discussion that follows with specific reference to the *ob* gene bears general applicability to the class of modulators that a part of the present invention, and is therefore to be accorded such latitude and scope of interpretation.

Therapeutic Implications

In the simplest analysis the *ob* gene determines body weight in mammals, in particular mice and man. The *ob* gene and, correspondingly, cognate molecules, appear to be part of a signaling pathway by which adipose tissue communicates

with the brain and the other organs. It is believed that the ob polypeptide is itself a signaling molecule, *i.e.*, a hormone. Alternatively ob may be responsible for the generation of a metabolic signal, *e.g.*, a stimulating hormone or an enzyme that catalyzes activation or synthesis of a peptide or steroid hormone. The most important piece of information for distinguishing between these possibilities or considering alternative hypothesis, is the complete DNA sequence of the RNA and its predicted protein sequence. Irrespective of its biochemical function the genetic data suggest that increased activity of *ob* would result in weight loss while decreased activity would be associated with weight gain. The means by which the activity of *ob* can be modified so as to lead to a therapeutic effect depends on its biochemical function.

Administration of recombinant ob polypeptide can result in weight loss. Recombinant protein can be prepared using standard bacterial and/or mammalian expression vectors, all as stated in detail earlier herein. Reduction of ob polypeptide activity (by developing antagonists, inhibitors, or antisense molecules) should result in weight gain as might be desirable for the treatment of the weight loss associated with cancer, AIDS or anorexia nervosa. Modulation of *ob* activity can be useful for reducing body weight (by increasing its activity) or increasing body weight (by decreasing its activity).

The ob polypeptide, or functionally active fragment thereof, or an antagonist thereof, can be administered orally or parenterally, preferably parenterally. Because metabolic homeostasis is a continuous process, controlled release administration of ob polypeptide is preferred. For example, the polypeptide may be administered using intravenous infusion, an implantable osmotic pump, a transdermal patch, liposomes, or other modes of administration. In one embodiment, a pump may be used (see Langer, supra; Sefton, *CRC Crit. Ref. Biomed. Eng.* 14:201 (1987); Buchwald et al., *Surgery* 88:507 (1980); Saudek et al., *N. Engl. J. Med.* 321:574 (1989)). In another embodiment, polymeric materials can be used (see *Medical Applications of Controlled Release*, Langer and

Wise (eds.), CRC Pres., Boca Raton, Florida (1974); *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, *J. Macromol. Sci. Rev. Macromol. Chem.* 23:61 (1983); see also Levy et al., *Science* 228:190 (1985); During et al., *Ann. Neurol.* 25:351 (1989); Howard et al., *J. Neurosurg.* 71:105 (1989)). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, i.e., the brain, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in *Medical Applications of Controlled Release*, supra, vol. 2, pp. 115-138 (1984)). Other controlled release systems are discussed in the review by Langer (*Science* 249:1527-1533 (1990)). In another embodiment, the therapeutic compound can be delivered in a vesicle, in particular a liposome (see Langer, *Science* 249:1527-1533 (1990); Treat et al., in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, ibid., pp. 317-327; see generally ibid.)

In a further aspect, recombinant cells that have been transformed with the *ob* gene and that express high levels of the polypeptide can be transplanted in a subject in need of ob polypeptide. Preferably autologous cells transformed with *ob* are transplanted to avoid rejection; alternatively, technology is available to shield non-autologous cells that produce soluble factors within a polymer matrix that prevents immune recognition and rejection.

Thus, the ob polypeptide can be delivered by intravenous, intraarterial, intraperitoneal, intramuscular, or subcutaneous routes of administration. Alternatively, the ob polypeptide, properly formulated, can be administered by nasal or oral administration. A constant supply of ob can be ensured by providing a therapeutically effective dose (*i.e.*, a dose effective to induce metabolic changes in a subject) at the necessary intervals, *e.g.*, daily, every 12 hours, etc. These parameters will depend on the severity of the disease condition being treated, other actions, such as diet modification, that are implemented, the weight, age, and sex

of the subject, and other criteria, which can be readily determined according to standard good medical practice by those of skill in the art.

Alternatively, the *ob* gene could be introduced into human fat cells to develop gene therapy for obesity. Such therapy would be expected to decrease body weight. Conversely, introduction of antisense constructs into human fat cells would reduce the levels of active ob polypeptide and would be predicted to increase body adiposity.

If ob is an enzyme, strategies have begun to be developed for the identification of the substrate and product of the catalyzed reaction that would make use of the recombinant protein. The rationale for this strategy is as follows: If *ob* is an enzyme that catalyzes a particular reaction in adipose tissue, then fat cells from *ob* mice should have high levels of the substrate and very little product. Since it is hypothesized that *db* mice are resistant to the product of this reaction, fat cells from *db* mice should have high levels of the reaction product. Thus, comparisons of lipid and peptide extracts of *ob* and *db* adipose tissue using gas chromatography or other chromatographic methods should allow the identification of the product and substrate of the key chemical reaction. The prediction would be that the recombinant ob protein would catalyze this reaction. The product of this reaction would then be a candidate for a signaling molecule that modulates body weight.

As noted above, the functional activity of the ob polypeptide can be effected transgenically,, *e.g.*, by gene therapy. In this respect, a transgenic mouse model can be used. The *ob* gene can be used in complementation studies employing transgenic mice. Transgenic vectors, including viral vectors, or cosmid clones (or phage clones) corresponding to the wild type locus of candidate gene, can be constructed using the isolated *ob* gene. Cosmids may be introduced into transgenic mice using published procedures (Jaenisch, *Science* **240**, 1468-1474, 1988). The constructs are introduced into fertilized eggs derived from an intercross between F1 progeny of a C57BL/6J *ob/ob* X DBA intercross. These

crosses require the use of C57BL/6J *ob/ob* ovarian transplants to generate the F1 animals. DBA/2J mice are used as the counterstrain because they have a nonagouti coat color which is important when using the ovarian transplants. Genotype at the *ob* loci in cosmid transgenic animals can be determined by typing animals with tightly linked RFLPs or microsatellites which flank the mutation and which are polymorphic between the progenitor strains. Complementation will be demonstrated when a particular construct renders a genetically obese F2 animal (as scored by RFLP analysis) lean and nondiabetic. Under these circumstances, final proof of complementation will require that the *ob/ob* or *db/db* animal carrying the transgene be mated to the *ob/ob* or *db/db* ovarian transplants. In this cross, all N2 animals which do not carry the transgene will be obese and insulin resistant/diabetic, while those that do carry the transgene will be lean and have normal glucose and insulin concentrations in plasma. In a genetic sense, the transgene acts as a suppressor mutation.

Alternatively, *ob* genes can be tested by examining their phenotypic effects when express in antisense orientation in wild-type animals. In this approach, expression of the wild type allele is suppressed, which leads to a mutant phenotype. RNA·RNAduplex formation (antisense·sense)prevents normal handling of mRNA, resulting in partial or complete elimination of wild-type gene effect. This technique has been used to inhibit Tk synthesis in tissue culture and to produce phenotypes of the Kruppel mutation in *Drosophila*, and the shiverer mutation in mice (Izant and Weintraub, *Cell* **36**, 1007-1015, 1984; Green et al., *Annu. Rev. Biochem.* **55**,569-597, 1986; Katsuki *et al., Science* **241**, 593-595, 1988). An important advantage of this approach is that only a small portion of the gene need be expressed for effective inhibition of expression of the entire cognate mRNA. The antisense transgene will be placed under control of its own promoter or another promoter expressed in the correct cell type, and placed upstream of the SV40 poly A site. This transgene will be used to make transgenic mice. Transgenic mice will also be mated ovarian transplants to test whether *ob* heterozygotes are more sensitive to the effects of the antisense construct.

In the long term, the elucidation of the biochemical function of the *ob* gene product (the ob polypeptide or protein) is useful for identifying small molecule agonists and antagonists that affect its activity.

Pharmaceutical Compositions

In yet another aspect of the present invention, provided are pharmaceutical compositions of the above. Such pharmaceutical compositions may be for administration for injection, or for oral, pulmonary, nasal or other forms of administration. In general, comprehended by the invention are pharmaceutical compositions comprising effective amounts of protein or derivative products of the invention together with pharmaceutically acceptable diluents, preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. Such compositions include diluents of various buffer content (e.g., Tris-HCl, acetate, phosphate), pH and ionic strength; additives such as detergents and solubilizing agents (e.g., Tween 80, Polysorbate 80), anti-oxidants (e.g., ascorbic acid, sodium metabisulfite), preservatives (e.g., Thimersol, benzyl alcohol) and bulking substances (e.g., lactose, mannitol); incorporation of the material into particulate preparations of polymeric compounds such as polylactic acid, polyglycolic acid, etc. or into liposomes. Hylauronic acid may also be used. Such compositions may influence the physical state, stability, rate of in vivo release, and rate of in vivo clearance of the present proteins and derivatives. See, e.g., Remington's Pharmaceutical Sciences, 18th Ed. (1990, Mack Publishing Co., Easton, PA 18042) pages 1435-1712 which are herein incorporated by reference. The compositions may be prepared in liquid form, or may be in dried powder, such as lyophilized form.

Oral Delivery

Contemplated for use herein are oral solid dosage forms, which are described generally in Remington's Pharmaceutical Sciences, 18th Ed.1990 (Mack Publishing Co. Easton PA 18042) at Chapter 89, which is herein incorporated by reference. Solid dosage forms include tablets, capsules, pills, troches or lozenges, cachets or pellets. Also, liposomal or proteinoid encapsulation may be used to formulate the present compositions (as, for example, proteinoid microspheres reported in U.S. Patent No. 4,925,673). Liposomal encapsulation may be used and the liposomes may be derivatized with various polymers (E.g., U.S. Patent No. 5,013,556). A description of possible solid dosage forms for the therapeutic

is given by Marshall, K. In: *Modern Pharmaceutics* Edited by G.S. Banker and C.T. Rhodes Chapter 10, 1979, herein incorporated by reference. In general, the formulation will include the protein (or chemically modified protein), and inert ingredients which allow for protection against the stomach environment, and release of the biologically active material in the intestine.

Also specifically contemplated are oral dosage forms of the above derivatized proteins. Protein may be chemically modified so that oral delivery of the derivative is efficacious. Generally, the chemical modification contemplated is the attachment of at least one moiety to the protein (or peptide) molecule itself, where said moiety permits (a) inhibition of proteolysis; and (b) uptake into the blood stream from the stomach or intestine. Also desired is the increase in overall stability of the protein and increase in circulation time in the body. Examples of such moieties include: Polyethylene glycol, copolymers of ethylene glycol and propylene glycol, carboxymethyl cellulose, dextran, polyvinyl alcohol, polyvinyl pyrrolidone and polyproline. Abuchowski and Davis, Soluble Polymer-Enzyme Adducts. In: "Enzymes as Drugs", Hocenberg and Roberts, eds., Wiley-Interscience, New York, NY, (1981), pp 367-383; Newmark, et al., J. Appl. Biochem. 4: 185-189 (1982). Other polymers that could be used are poly-1,3-dioxolane and poly-1,3,6-tioxocane. Preferred for pharmaceutical usage, as indicated above, are polyethylene glycol moieties.

For the protein (or derivative) the location of release may be the stomach, the small intestine (the duodenum, the jejunem, or the ileum), or the large intestine. One skilled in the art has available formulations which will not dissolve in the stomach, yet will release the material in the duodenum or elsewhere in the intestine. Preferably, the release will avoid the deleterious effects of the stomach environment, either by protection of the protein (or derivative) or by release of the biologically active material beyond the stomach environment, such as in the intestine.

To ensure full gastric resistance a coating impermeable to at least pH 5.0 is essential. Examples of the more common inert ingredients that are used as enteric coatings are cellulose acetate trimellitate (CAT), hydroxypropylmethylcellulose phthalate (HPMCP), HPMCP 50, HPMCP 55, polyvinyl acetate phthalate (PVAP), Eudragit L30D, Aquateric, cellulose acetate phthalate (CAP), Eudragit L, Eudragit S, and Shellac. These coatings may be used as mixed films.

A coating or mixture of coatings can also be used on tablets, which are not intended for protection against the stomach. This can include sugar coatings, or coatings which make the tablet easier to swallow. Capsules may consist of a hard shell (such as gelatin) for delivery of dry therapeutic i.e. powder; for liquid forms, a soft gelatin shell may be used. The shell material of cachets could be thick starch or other edible paper. For pills, lozenges, molded tablets or tablet triturates, moist massing techniques can be used.

The therapeutic can be included in the formulation as fine multiparticulates in the form of granules or pellets of particle size about 1mm. The formulation of the material for capsule administration could also be as a powder, lightly compressed plugs or even as tablets. The therapeutic could be prepared by compression.

Colorants and flavoring agents may all be included. For example, the protein (or derivative) may be formulated (such as by liposome or microsphere encapsulation) and then further contained within an edible product, such as a refrigerated beverage containing colorants and flavoring agents.

One may dilute or increase the volume of the therapeutic with an inert material. These diluents could include carbohydrates, especially mannitol, α-lactose, anhydrous lactose, cellulose, sucrose, modified dextrans and starch. Certain inorganic salts may be also be used as fillers including calcium triphosphate, magnesium carbonate and sodium chloride. Some commercially available diluents are Fast-Flo, Emdex, STA-Rx 1500, Emcompress and Avicell.

Disintegrants may be included in the formulation of the therapeutic into a solid dosage form. Materials used as disintegrates include but are not limited to starch including the commercial disintegrant based on starch, Explotab. Sodium starch glycolate, Amberlite, sodium carboxymethylcellulose, ultramylopectin, sodium alginate, gelatin, orange peel, acid carboxymethyl cellulose, natural sponge and bentonite may all be used. Another form of the disintegrants are the insoluble cationic exchange resins. Powdered gums may be used as disintegrants and as binders and these can include powdered gums such as agar, Karaya or tragacanth. Alginic acid and its sodium salt are also useful as disintegrants.

Binders may be used to hold the therapeutic agent together to form a hard tablet and include materials from natural products such as acacia, tragacanth, starch and gelatin. Others include methyl cellulose (MC), ethyl cellulose (EC) and carboxymethyl cellulose (CMC). Polyvinyl pyrrolidone (PVP) and hydroxypropylmethyl cellulose (HPMC) could both be used in alcoholic solutions to granulate the therapeutic.

An antifrictional agent may be included in the formulation of the therapeutic to prevent sticking during the formulation process. Lubricants may be used as a layer between the therapeutic and the die wall, and these can include but are not limited to; stearic acid including its magnesium and calcium salts, polytetrafluoroethylene (PTFE), liquid paraffin, vegetable oils and waxes. Soluble lubricants may also be used such as sodium lauryl sulfate, magnesium lauryl sulfate, polyethylene glycol of various molecular weights, Carbowax 4000 and 6000.

Glidants that might improve the flow properties of the drug during formulation and to aid rearrangement during compression might be added. The glidants may include starch, talc, pyrogenic silica and hydrated silicoaluminate.

To aid dissolution of the therapeutic into the aqueous environment a surfactant might be added as a wetting agent. Surfactants may include anionic detergents such as sodium lauryl sulfate, dioctyl sodium sulfosuccinate and dioctyl sodium sulfonate. Cationic detergents might be used and could include benzalkonium chloride or benzethomium chloride. The list of potential nonionic detergents that could be included in the formulation as surfactants are lauromacrogol 400, polyoxyl 40 stearate, polyoxyethylene hydrogenated castor oil 10, 50 and 60, glycerol monostearate, polysorbate 40, 60, 65 and 80, sucrose fatty acid ester, methyl cellulose and carboxymethyl cellulose. These surfactants could be present in the formulation of the protein or derivative either alone or as a mixture in different ratios.

Additives which potentially enhance uptake of the protein (or derivative) are for instance the fatty acids oleic acid, linoleic acid and linolenic acid.

Controlled release formulation may be desirable. The drug could be incorporated into an inert matrix which permits release by either diffusion or leaching mechanisms i.e. gums. Slowly degenerating matrices may also be incorporated into the formulation. Another form of a controlled release of this therapeutic is by a method based on the Oros therapeutic system (Alza Corp.), i.e. the drug is enclosed in a semipermeable membrane which allows water to enter and push drug out through a single small opening due to osmotic effects. Some entric coatings also have a delayed release effect.

Other coatings may be used for the formulation. These include a variety of sugars which could be applied in a coating pan. The therapeutic agent could also be given in a film coated tablet and the materials used in this instance are divided into 2 groups. The first are the nonenteric materials and include methyl cellulose, ethyl cellulose, hydroxyethyl cellulose, methylhydroxy-ethyl cellulose, hydroxypropyl cellulose, hydroxypropyl-methyl cellulose, sodium carboxy-methyl

cellulose, providone and the polyethylene glycols. The second group consists of the enteric materials that are commonly esters of phthalic acid.

A mix of materials might be used to provide the optimum film coating. Film coating may be carried out in a pan coater or in a fluidized bed or by compression coating.

Pulmonary Delivery

Also contemplated herein is pulmonary delivery of the present protein (or derivatives thereof). The protein (or derivative) is delivered to the lungs of a mammal while inhaling and traverses across the lung epithelial lining to the blood stream. (Other reports of this include Adjei et al., PHARMACEUTICAL RESEARCH, VOL. 7, No. 6, pp. 565-569 (1990); Adjei et al., *International Journal of Pharmaceutics*, Vol. 63, pp. 135-144 (1990)(leuprolide acetate); Braquet et al., *Journal of Cardiovascular Pharmacology*, Vol. 13, suppl. 5, s. 143-146 (1989)(endothelin-1);Hubbard et al., *Annals of Internal Medicine*, Vol. III, No. 3, pp. 206-212(1989)(α1- antitrypsin); Smith et al., *J. Clin. Invest.*, Vol. 84, pp. 1145-1146 (1989)(α1-proteinase); Oswein et al., "*Aerosolization of Proteins*", Proceedings of Symposium on Respiratory Drug Delivery II, Keystone, Colorado, March, 1990 (recombinant human growth hormone); Debs et al., *The Journal of Immunology*, Vol. 140, pp. 3482-3488 (1988)(interferon-γ and tumor necrosis factor alpha) and Platz et al., U.S. Patent No. 5,284,656 (granulocyte colony stimulating factor). Contemplated for use in the practice of this invention are a wide range of mechanical devices designed for pulmonary delivery of therapeutic products, including but not limited to nebulizers, metered dose inhalers, and powder inhalers, all of which are familiar to those skilled in the art.

Some specific examples of commercially available devices suitable for the practice of this invention are the Ultravent nebulizer, manufactured by Mallinckrodt, Inc., St. Louis, Missouri; the Acorn II nebulizer, manufactured by Marquest Medical Products, Englewood, Colorado; the Ventolin metered dose inhaler, manufactured

by Glaxo Inc., Research Triangle Park, North Carolina; and the Spinhaler powder inhaler, manufactured by Fisons Corp., Bedford, Massachusetts.

All such devices require the use of formulations suitable for the dispensing of protein (or derivative). Typically, each formulation is specific to the type of device employed and may involve the use of an appropriate propellant material, in addition to the usual diluents, adjuvants and/or carriers useful in therapy. Also, the use of liposomes, microcapsules or microspheres, inclusion complexes, or other types of carriers is contemplated. Chemically modified protein may also be prepared in different formulations depending on the type of chemical modification or the type of device employed.

Formulations suitable for use with a nebulizer, either jet or ultrasonic, will typically comprise protein (or derivative) dissolved in water at a concentration of about 0.1 to 25 mg of biologically active protein per mL of solution. The formulation may also include a buffer and a simple sugar (e.g., for protein stabilization and regulation of osmotic pressure). The nebulizer formulation may also contain a surfactant, to reduce or prevent surface induced aggregation of the protein caused by atomization of the solution in forming the aerosol.

Formulations for use with a metered-dose inhaler device will generally comprise a finely divided powder containing the protein (or derivative) suspended in a propellant with the aid of a surfactant. The propellant may be any conventional material employed for this purpose, such as a chlorofluorocarbon, a hydrochlorofluorocarbon, a hydrofluorocarbon, or a hydrocarbon, including trichlorofluoromethane, dichlorodifluoromethane, dichlorotetrafluoroethanol, and 1,1,1,2-tetrafluoroethane, or combinations thereof. Suitable surfactants include sorbitan trioleate and soya lecithin. Oleic acid may also be useful as a surfactant.

Formulations for dispensing from a powder inhaler device will comprise a finely divided dry powder containing protein (or derivative) and may also include a

bulking agent, such as lactose, sorbitol, sucrose, or mannitol in amounts which facilitate dispersal of the powder from the device, e.g., 50 to 90% by weight of the formulation. The protein (or derivative) should most advantageously be prepared in particulate form with an average particle size of less than 10 μm (or microns), most preferably 0.5 to 5 μm, for most effective delivery to the distal lung.

Nasal Delivery

Nasal delivery of the protein (or derivative) is also contemplated. Nasal delivery allows the passage of the protein to the blood stream directly after administering the therapeutic product to the nose, without the necessity for deposition of the product in the lung. Formulations for nasal delivery include those with dextran or cyclodextran.

Methods of Treatment, Methods of Preparing a Medicament

In yet another aspect of the present invention, methods of treatment and manufacture of a medicament are provided. Conditions alleviated or modulated by the administration of the present derivatives are those indicated above.

Dosages

For all of the above molecules, as further studies are conducted, information will emerge regarding appropriate dosage levels for treatment of various conditions in various patients, and the ordinary skilled worker, considering the therapeutic context, age and general health of the recipient, will be able to ascertain proper dosing. Generally, for injection or infusion, dosage will be between 0.01 μg of biologically active protein/kg body weight, (calculating the mass of the protein alone, without chemical modification), and 100 μg/kg (based on the same). The dosing schedule may vary, depending on the circulation half-life of the protein or derivative used, and the formulation used.

Administration with other compounds

For therapy associated with obesity, one may administer the present protein (or derivatives) in conjunction with one or more pharmaceutical compositions used for treating other clinical complications of obesity, such as those used for treatment of diabetes (e.g., insulin), high blood pressure, high cholesterol, and other adverse conditions incident to obesity. Also, other appetite suppressants may be co-administered, e.g., amphetamines. Administration may be simultaneous (for example, administration of a mixture of the present protein and insulin) or may be in serriatim.

Diagnostic Implications

The human cDNA clones that have recently been isolated have been sequenced as presented herein. This facilitates the determination of the complete sequence of the human gene (see Figure 20). DNA sequences from the introns of the human *ob* gene have been obtained (Figure 20), and these have been used to prepare PCR primers to PCR amplify the coding sequence of the *ob* gene from human genomic DNA so as to identify mutations or allelic variants of the *ob* gene, all in accordance with protocols described in detail earlier herein. Specific PCR primers for amplifying human genomic *ob* are described in a specific Example, *infra*.

The current hypothesis is that heterozygous mutations in the *ob* gene will be associated with mild/moderate obesity while homozygous mutations would be associated with several DNA sequence based diagnostic tests obesity. If this is true, it would allow the ascertainment of people at risk for the development of obesity and make possible the application of drug treatment and/or lifestyle changes before an increased body weight is fully developed.

Alternatively, the presence of microsatellites that segregate with mutant forms of human *ob* can be used for diagnosis. Various PCR primers, including those based on the nucleotide sequence provided in Figure 20A, can be used in this respect.

The *ob* gene may also be useful diagnostically for measurements of its encoded RNA and protein in nutritional disorders. It will be of importance to know, in a particular nutritional disorder, whether *ob* RNA and/or protein is unregulated or downregulated. Thus, if an obese person has increased levels of *ob*, it would appear that the problem is downstream of *ob*, while if *ob* is reduced, it would appear that inappropriately low levels of *ob* may be cause of obesity (whether or not the defect is in the *ob* gene). Conversely, if a cancer or AIDS patient who lost weight had elevated levels of *ob*, it may be concluded that inappropriately high expression of *ob* is responsible for the weight loss.

The cloned human cDNA will be of use for the measurement of the levels of human *ob* RNA. In addition, recombinant human protein will be prepared and used to develop immunoassays to enable measurement of the fat and perhaps plasma levels of the ob protein.

<u>Agricultural Applications</u>

The *ob* gene can also be isolated from domestic animals, and the corresponding ob polypeptide obtained thereby. In a specific example, *infra*, the a probe derived from the murine *ob* gene hybridizes to corresponding homologous coding sequences from a large number of species of animals. As discussed for human therapies, recombinant proteins can also be prepared and administered to domestic animals. Administration of the polypeptide can be implemented to produce leaner food animals, such as beef cattle, swine, poultry, sheep, etc. Preferably, an autologous ob polypeptide is administered, although the invention contemplates administration of anti-autologous polypeptide as well. Since the ob polypeptide consists of approximately 160 amino acid residues, it may not be highly immunogenic. Thus, administration of non-autologous polypeptide may not result in an immune response.

Alternatively, the introduction of the cloned genes into transgenic domestic animals would allow one to potentially decrease body weight and adiposity by

overexpressing an *ob* transgene. The simplest means of achieving this would be to target an *ob* transgene to fat using its own or another fat specific promoter.

Conversely, increases in body fat might be desirable in other circumstances such as for the development of Kobe beef or fatty liver to make foie gras. This could be accomplished by targeting an antisense *ob* transgene to fat, or by using gene knockout technology. Alternatively, where an increase in body weight at percentage of fat is desired, an inhibitor or antagonist of the ob polypeptide can be administered. Such inhibitors or antagonists include, but are not limited to, antibodies reactive with the polypeptide, and fragments of the polypeptide that bind but do not activate the ob receptor, *i.e.*, antagonists of ob polypeptide.

<u>The ob Receptor</u>

Development of small molecule agonists and antagonists of the *ob* factor will be greatly facilitated by the isolation of its receptor. This can be accomplished by preparing active ob polypeptide and using it to screen an expression library using standard methodology. Receptor binding in the expression library can be tested by administering recombinant polypeptide prepared using either bacterial or mammalian expression vectors, and observing the effects of short term and continuous administration of the recombinant polypeptide on the cells of the expression library, or by directly detecting binding of ob polypeptide to the cells.

As it is presently believed that the ob receptor is likely to be located in the hypothalamus and perhaps liver, preferably cDNA libraries from these tissues will be constructed in standard expression cloning vectors. These cDNA clones would next be introduced into COS cells as pools and the resulting transformants would be screened with active ligand to identify COS cells expressing the ob receptor. Positive clones can then be isolated so as to recover the cloned receptor. The cloned receptor would be used in conjunction with the *ob* ligand (assuming it is a hormone) to develop the necessary components for screening of small molecule modulators of *ob*.

Once a recombinant which expresses the ob receptor gene sequence is identified, the recombinant *ob* receptor can be analyzed. This is achieved by assays based on the physical or functional properties of the *ob* receptor, including radioactive labelling of the receptor followed by analysis by gel electrophoresis, immunoassay, ligand binding, etc. Furthermore, antibodies to the ob receptor could be generated as described above.

The structure of the *ob* receptor can be analyzed by various methods known in the art. Preferably, the structure of the various domains, particularly the *ob* binding site, is analyzed. Structural analysis can be performed by identifying sequence similarity with other known proteins, particular hormone and protein receptors. The degree of similarity (or homology) can provide a basis for predicting structure and function of the *ob* receptor, or a domain thereof. In a specific embodiment, sequence comparisons can be performed with sequences found in GenBank, using, for example, the FASTA and FASTP programs (Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85:2444-48).

The protein sequence can be further characterized by a hydrophilicity analysis (*e.g.*, Hopp and Woods, 1981, Proc. Natl. Acad. Sci. U.S.A. 78:3824). A hydrophilicity profile can be used to identify the hydrophobic and hydrophilic regions of the *ob* receptor protein, which may in turn indicate extracytoplasmic, membrane binding, and intracytoplasmic regions.

Secondary structural analysis (*e.g.*, Chou and Fasman, 1974, Biochemistry 13:222) can also be done, to identify regions of the *ob* receptor that assume specific secondary structures.

Manipulation, translation, and secondary structure prediction, as well as open reading frame prediction and plotting, can also be accomplished using computer software programs available in the art.

By providing an abundant source of recombinant *ob* polypeptide, and the opportunity to isolate the *ob* receptor (*i.e.*, the *db* gene product), the present invention enables quantitative structural determination of the active conformation of the *ob* polypeptide and the *ob* receptor, or domains thereof. In particular, enough material is provided for nuclear magnetic resonance (NMR), infrared (IR), Raman, and ultraviolet (UV), especially circular dichroism (CD), spectroscopic analysis. In particular NMR provides very powerful structural analysis of molecules in solution, which more closely approximates their native environment (Marion et al., 1983, Biochem. Biophys. Res. Comm. 113:967-974; Bar et al., 1985, J. Magn. Reson. 65:355-360; Kimura et al., 1980, Proc. Natl. Acad. Sci. U.S.A. 77:1681-1685). Other methods of structural analysis can also be employed. These include but are not limited to X-ray crystallography (Engstom, A., 1974, Biochem. Exp. Biol. 11:7-13).

More preferably, co-crystals of ob polypeptide and ob receptor can be studied. Analysis of co-crystals provides detailed information about binding, which in turn allows for rational design of ligand agonists and antagonists. Computer modeling can also be used, especially in connection with NMR or X-ray methods (Fletterick, R. and Zoller, M. (eds.), 1986, Computer Graphics and Molecular Modeling, in *Current Communications in Molecular Biology*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

Identification and isolation of a gene encoding an ob receptor of the invention provides for expression of the receptor in quantities greater than can be isolated from natural sources, or in indicator cells that are specially engineered to indicate the activity of a receptor expressed after transfection or transformation of the cells. According, in addition to rational design of agonists and antagonists based on the structure of ob polypeptide, the present invention contemplates an alternative method for identifying specific ligands of ob receptor using various screening assays known in the art.

Any screening technique known in the art can be used to screen for *ob* receptor agonists or antagonists. The present invention contemplates screens for small molecule ligands or ligand analogs and mimics, as well as screens for natural ligands that bind to and agonize or antagonize activates *ob* receptor *in vivo*.

Knowledge of the primary sequence of the receptor, and the similarity of that sequence with proteins of known function, can provide an initial clue as the inhibitors or antagonists of the protein. Identification and screening of antagonists is further facilitated by determining structural features of the protein, *e.g.*, using X-ray crystallography, neutron diffraction, nuclear magnetic resonance spectrometry, and other techniques for structure determination. These techniques provide for the rational design or identification of agonists and antagonists.

Another approach uses recombinant bacteriophage to produce large libraries. Using the "phage method" (Scott and Smith, 1990, Science 249:386-390; Cwirla, et al., 1990, Proc. Natl. Acad. Sci., 87:6378-6382; Devlin et al., 1990, Science, 249:404-406), very large libraries can be constructed (10^6-10^8 chemical entities). A second approach uses primarily chemical methods, of which the Geysen method (Geysen et al., 1986, Molecular Immunology 23:709-715; Geysen et al. 1987, J. Immunologic Method 102:259-274) and the recent method of Fodor et al. (1991, Science 251, 767-773) are examples. Furka et al. (1988, 14th International Congress of Biochemistry, Volume 5, Abstract FR:013; Furka, 1991, Int. J. Peptide Protein Res. 37:487-493), Houghton (U.S. Patent No. 4,631,211, issued December 1986) and Rutter et al. (U.S. Patent No. 5,010,175, issued April 23, 1991) describe methods to produce a mixture of peptides that can be tested as agonists or antagonists.

In another aspect, synthetic libraries (Needels et al., 1993, "Generation and screening of an oligonucleotide encoded synthetic peptide library," Proc. Natl. Acad. Sci. USA 90:10700-4; Lam et al., International Patent Publication No. WO 92/00252, each of which is incorporated herein by reference in its entirety), and

the like can be used to screen for *ob* receptor ligands according to the present invention. With such libraries, receptor antagonists can be detected using cell that express the receptor without actually cloning the *ob* receptor (Lam et al., *supra*).

Alternatively, assays for binding of soluble ligand to cells that express recombinant forms of the ob receptor ligand binding domain can be performed. The soluble ligands can be provided readily as recombinant or synthetic ob polypeptide.

The screening can be performed with recombinant cells that express the ob receptor, or alternatively, using purified receptor protein, *e.g.*, produced recombinantly, as described above. For example, the ability of labeled, soluble or solubilized ob receptor that includes the ligand-binding portion of the molecule, to bind ligand can be used to screen libraries, as described in the foregoing references.

EXAMPLE SECTION

The following outlines the method used to identify the genetic material that is exemplary of the present invention. This endeavor comprises four sequential steps: A) Genetic Mapping, B) Physical Mapping, C) Candidate Gene Isolation, and D) Mutation detection. Following confirmation that the murine gene in object was isolated (Step D), the homologous human gene was sought, and both the murine and human genes and putative proteins were characterized. The steps are summarized in greater detail, below.

A. Genetic Mapping

The *ob* mutation was segregated in genetic crosses and standard linkage analysis was used to position the mutation relative to RFLPs (restriction fragment length polymorphisms). These data placed the *ob* gene in an ~ 5cM interval on proximal mouse chromosome 6. (5cM is a measurement of genetic distance

corresponding to 5 apparent genetic crossovers per 100 animals.) A total of 771 informative meioses were generated and used in subsequent genetic mapping (Friedman et al. *Genomics* **11**: 1054-1062, 1991). The genetic loci that were mapped relative to *ob* were all previously published. The two closest RFLPs described were defined by probes derived from the carboxypeptidase and *met* oncogene genes.

The genetic resolution of the experiments described above was inadequate to clone *ob*, principally because none of the genetic markers were in tight linkage. In order to identify the requisite tightly linked RFLPs, additional probes were isolated and the genetic cross was expanded. A method known as chromosome microdissection was used to isolate random pieces of DNA from proximal mouse chromosome 6 (Bahary et al., *Mammalian Genome* **4**: 511-515, 1993). Individual cloned probes were tested for tight linkage to *ob*. On the basis of these studies one probe, D6Rck13, also termed psd3, was selected for further analysis owing to its genetic proximity to *ob*.

This probe was used to genotype 835 *ob* progeny from interspecific and intersubspecific crosses, which indicated that D6Rck13 is nonrecombinant in all 835 animals as reported in Bahary et al. In the course of physical mapping, a new polymorphic marker was identified from a cosmid subclone derived from YAC 53A6. This new marker was positioned between D6Rck13 and the *ob* gene and was used to genotype the additional 771 informative meioses from intraspecific intercross and backcross. A single animal #167 was identified to bear a recombination crossover between *ob* and D6Rck39. These studies indicated that D6Rck39/D6RcK13 is ~0.06 cM from *ob*. An additional probe, Pax-4, was identified that was .12 cM proximal to *ob*. Pax-4 was recombinant in two animals; #111 and 420. Pax-4 is a pseudogene that was previously mapped to proximal mouse chromosome 6 by Gruss and co-workers (Gruss et al. *Genomics* **11**:424-434, 1991). On this basis, it was determined that the *ob* gene resides in

the ~ 0.2cM interval between Pax-4 and D6Rck13. This led to efforts to clone the interposing DNA in an effort to isolate *ob*.

B. Physical Mapping

The cloning of the DNA in this interval made use of yeast artificial chromosomes (YACs), a relatively new cloning vector that allows the cloning of long stretches of contiguous DNA often more than 1 million base pairs in length.

Firstly, yeast artificial chromosomes were isolated using D6Rck13 and Pax-4. This was accomplished by preparing purified DNA probes and using them to isolate the corresponding YACs. These YACs (#8, 16, 107 and 24) were isolated and initially characterized, and on the basis of the resulting analyses it was concluded that YAC 16 was the YAC that extended furthest distally, *i.e.*, closest to *ob*. The key end of YAC #16 was then recovered, and it was determined that this end was closer to *ob* than Pax-4. This end was termed 16M(+). This conclusion was reached because it was shown that this probe was not recombinant in animal #420 (as was Pax-4). This end was sequenced and used to develop a PCR assay. This PCR assay was used to screen a YAC library. Four positive clones were isolated. Subsequent characterization of these YACs by end-rescuing, restriction mapping, pulse field gel electrophoresis, and Southern blots with the genetic crosses determined that two of these YACs, *adu* and *aad*, were critical for subsequent studies. YAC *aad* is a 550 kB nonchimeric YAC which extended furthest distally. Therefore, the distal end of this YAC, *aad*(pICL) was used to complete the physical map. YAC *adu* is 370 kB nonchimeric YAC and its distal end, *adu*(+), was determined to be nonrecombinant in all the *ob* progeny of the genetic crosses including animals #111 and 167, suggesting that the *ob* gene might reside in this YAC.

A PCR assay for these two ends, *aad*(pICL) and *adu*(+) was developed and used for isolating moire YACs and P1 clones to continue physical mapping. The important P1 clones isolated by this effort included 498, 499, 500 (isolated using a

probe derived from *aad*(pICL)) and 322, 323 and 324 (using a probe from *adu*(+)).

In the meantime, YACs isolated by D6Rck13 (53A6, 25A8, 25A9, 25A10) were characterized. These studies determined that 53A6 extended furthest proximally toward the *aad* YAC. The size of the gap between 53A6 and *aad* was determined ~70 kB. The key end of 53A6, 53(pICL) was then used to screen three available YAC libraries and a P1 library. A critical P1 clone, 325, was isolated. This P1 clone overlapped with the P1 clones isolated by *aad*(pICL) as described above, and therefore served to close the gap between 53(pICL) and *aad*(pICL). As a result, the whole contig, containing YACs and P1 clones, of ~2.5 million base pairs in length, and which spanned Pax4, 16M(+), *adu*(+), *aad*(pICL), 53(pICL), D6Rck39 and D6Rck13, was cloned. By carefully mapping the sites of recombination apparent in animal #111 and 167, it was concluded that *ob* was situated in a 400 kB interval. To provide a working DNA source for isolating the *ob* gene, about 500 kB covering this nonrecombination region was isolated in a total of 24 P1 clones. These P1 clones, including 322 and 323, which later were proved to be useful clones, were used for exon trapping.

The physical map of the portion of the chromosome carrying *ob* is shown in Figure 7A. Figure 7B represents the YAC contig. Figure 7C represents the P1 contig.

C. Isolation of Candidate Genes

The method used to isolate genes in this interval was exon trapping. This method used a commercial vector to identify exon DNA (*i.e.*, coding sequences) by selecting for functional splice acceptor and donor sequences in genomic DNA introduced into a test construct. The DNA from these P1s were grown and subcloned into the exon trapping vector. These clones were short inserts cloned into a Bluescript vector. Each clone was PCR amplified with PCR primers corresponding to plasmid sequences that flanked the insert. The PCR

amplification was performed directly on the bacteria that carried the plasmid. The reactions were set up using a Biomek robot. The PCR products were electrophoresed on a 1% agarose gel in TBE buffer that contained ethidium bromide. The exon trapping technique was modified to eliminate contaminating *E. coli* DNA from the P1 clones, and to screen out the abundant artifactual exons, which exceeded 80-90% of the putative exons trapped. The exon trapping vector includes HIV sequences; a short segment of these vector sequences corresponds to this artifact.

The exon trapping experiment was performed using various P1 clones. Exon trapping products were then amplified by PCR, selected, and sequenced. Sequences of putative "exons" were compared with those in Genbank using the Blast computer program. About 15 exons were selected for further examination by RT-PCR, Northern analysis, and zoo blot for the presence of corresponding RNA or conservative sequences. Seven of the 15 putative exons, 325-2, 323-9, 322-5, D1-F7, 1H3, and 2G7, were found to encode an RNA transcript. 325-2 is a testis specific gene; 323-8 and 323-9 are likely two exons from the same gene expressed mainly in brain and kidney. IH3 and 322-5 represent two low level brain transcripts. D1-F7 is an exon from a previously cloned gene, inosine monophosphate dehydrogenase (IMPDH), which has ubiquitous expression pattern. None of these genes appeared to encode *ob*. 2G7, which is the *ob* exon, is discussed further below.

After three unsuccessful efforts to exon trap the *ob* gene, another attempt was made by pooling DNA from all the P1s from the critical *ob* region. These included P1s: 258, 259, 322, 323, 324, 325, 498, 499, 500, 653, 654 and others. Thereafter P1s 258, 260, 322, 498 and 499 were subcloned into the exon trapping vector, and subsequently several plates were prepared with bacterial clones, each of which carried a putative exon. Approximately 192 clones representing putative *ob* candidates were obtained. As noted above, a consistent artifact such that many of the isolates contained two trapped exons derived from the vector was observed.

Thus, clones were identified both by their size and the fact that hybridization of DNA probes corresponding to this artifact hybridized to the corresponding bands on a Southern blot of the gel. In this way, 185 out of 192 clones were excluded from further evaluation. Exclusion of the artifacts on the basis of size alone was not possible, as this could have, in the end, led to exclusion of the exon corresponding to *ob*.

Thus, of the 192 exons, a total of seven exons were selected for further study. Templates for sequencing the seven exons were prepared, and sequencing was performed. The sequences for the 7 exons were analyzed and it was found that 4 were identical and one was an apparent artifact. In particular, clone 1D12 contained the "HIV sequence," *i.e.*, the artifact band. This left three exons for further analysis: 1F1, 2G7 and 1H3. 1F1 was eliminated because it mapped outside the critical region. PCR primers for both 1H3 and 2G7 were selected and synthesized.

The sequence of the exon on 2G7 was determined, and is shown in Figure 10 (SEQ ID NO:7). PCR primers for 2G7 were selected and synthesized. The portions of the sequence corresponding to the PCR primers are underlined. The primers used were:

5′ CCA GGG CAG GAA AAT GTG (Tm = 60.0) (SEQ ID NO:8)

3′ CAT CCT GGA CTT TCT GGA TAG G (Tm = 60.0) (SEQ ID NO:9)

These primers amplified genome DNA with PCR conditions as follows: 25-30 cycles at 55° annealing x 2', 72° extension x 2', 94° denaturation x 1' in standard PCR buffer. These primers were also used to generate a labeled probe by including ^{32}P dCTP in the PCR reaction with a corresponding reduction in the amount of cold dCTP.

A RT PCR was performed on a variety of tissue RNAs and it was concluded that 2G7 was expressed exclusively in white fat among the tissues examined (Figure 11A). Thereafter, ^{32}P-labelled 2G7 was hybridized to a Northern blot of tissue RNAs (Figure 11B) and showed that its RNA was expressed at high level in fat tissue but was either not expressed or expressed at very low levels in all other tissues (where the signals may be the result of fat contaminating the tissue preparations). Ten μg of total RNA from each of the tissues listed was electrophoresed on an agarose gel with formaldehyde. The probe was hybridized at 65° in a standard hybridization buffer, Rapid Hype (Amersham). The size of the RNA was approximately 4.9 kB. At this point 2G7 was considered to be a viable candidate gene for *ob* and was analyzed further.

D. Mutation Detection

In order to confirm that 2G7 encoded the *ob* gene, it was necessary to demonstrate differences in the levels of RNA expression of DNA sequence of this gene in mutant as compared to wild type animals. Two separate mutations of the *ob* gene are available for study, C57BL/6J ob/ob (1J) and Ckc/Smj ob/ob (2J). These will be referred hereinafter as 1J and 2J, respectively. (Informal nomenclature is used to refer to the mouse strains studied. Throughout this specification and in the drawings, it will be understood that C57BL/6J refers to C57BL/6J +/+; CKC/smj refers to SM/Ckc-+^{Dac}-+/+; CKC/smj ob/ob refers to SM/Ckc-+^{Dac}-ob^{2J}/ob^{2J}). RNA was prepared from fat tissue that had been isolated from 1J, 2J, and control animals. Total RNA for each sample was treated with DNase and then reverse transcribed using oligo-dT as a primer and reverse transcriptase. The resulting single stranded cDNA was then PCR amplified either with the 2G7 primers (conditions shown above) for the lower band or commercially available actin primers for the upper band. The RT PCR products were run on a 1% agarose TBE gel that was stained with ethidium bromide (Figure 12A). Using RT-PCT it was found that while 2G7 mRNA was expressed in 1J and all the other control mice, it was completely missing in 2J mouse. No signal was detected after 30

cycles of amplification. This experiment provided direct evidence that 2G7 corresponded to an exon from the *ob* gene.

Since 2J mutation is relatively recent and is maintained as a coisogenic strain, this result was the first available evidence that indicated that 2G7 is an exon from the *ob* gene. The mutation is likely located in the promoter region which leads to total abortion of the mRNA synthesis. The presence of signal in 1J mouse in this RT-PCT experiment suggested that 1J might carry a point mutation which does not result in a gross change in size of the RNA sample. In addition, 2G7 mRNA was absent, when tested by RT PCR, from four additional 2J animals.

This result was confirmed on a Northern blot (Figure 12B). Fat cell RNA was prepared from each of the strains (C57Bl/6J, 1J, CKC/smj, and 2J). Ten μg of these RNAs were run out. The blot was probed with the 2G7 probe that was PCR labeled, by amplification of the material, i.e., band, in Figure 11 using ^{32}P-dCTP in the PCR reaction. Actin is a control for the amount of RNA loaded. The actin signal is fairly similar in all of the samples. The *ob* signal is absent in brain because the mRNA is specific to fat cells.

The results of the Northern analysis confirm that 2G7 RNA was absent in 2J mice. The *ob* RNA is absent in the CKC/smj ob/ob mice because in this obese mutant strain the gene is disrupted such that no RNA is made. In addition, the level of 2G7 RNA was increased ~10-20 fold in 1J as well as db/db fat. These results are compatible with the hypothesis that *ob* either encodes circulating hormone or is responsible for the generation of a signal from fat cells that modulates body weight. These results supported the conclusion that 2G7 is the *ob* gene and predicted that 1J mice have a point mutation, probably a nonsense mutation leading to a premature translation termination.

These Northern results have been replicated using fat cell RNA preparations from four different 2J animals (Figure 13). In this assay, *ap2* is a fat-specific transcript

that was used as a control much the same as actin in Figure 12B. There is no significance to the varying density of the ap2 band. ap2 was labeled by designing PCR primers form the published ap2 sequence. The RT PCR products of fat cell RNA were then relabeled using the same protocol for PCR labeling. This analysis demonstrates the presence of *ob* mRNA in normal homozygous or heterozygous animals, and its absence from 2J mutant animals.

The mutation has been identified in 1J mice. The mutation is a C to T base change that results in change of an arginine to an apparent premature stop codon at amino acid 108, and in all likelihood accounts for the 1J mutation (Figure 14) despite high level expression of the *ob* mRNA (see Figure 12 and 13, C57BL/6J ob/ob lanes).

More recently, Southern blots have been used to conclude that the 2J mutation is the result of a detectable DNA change at the 5′ end of *ob* that appears to completely abolish RNA expression. The exact nature of this possible rearrangement remains to be determined.

A genomic Southern blot of DNA from the CKC/smj (SM/Ckc-+^{Dac}) and C57BL/6J mice using four different restriction endonucleases was performed in order to determine whether the mutant *ob* yielded a unique fragment pattern (Figure 15A). Approximately 10 μg of DNA (derived from genomic DNA prepared from liver, kidney, or spleen) was restriction digested with the restriction enzyme indicated. The DNA was then electrophoresed in a 1% agarose TBE gel. The DNA was transferred to an imobilon membrane and hybridized to the PCR labeled 2G7 probe. The key band is the uppermost band in the BglII digest for the CKC/smj ob/ob (SM/Ckc-+^{DAC} ob^{2J}/ob^{2J}) DNA. This band is of higher molecular weight than in the other strain, indicating a mutation in this strain.

Figure 15B is a southern blot of a BglII digest of genomic DNA from the progeny of an ob^{2J}/+ x ob^{2J}/+ cross. Some of the DNAs have only the upper band, some

only the lower band, and some have the both bands. The animals with only the upper band are allo-obese, *i.e.*, ob^{2J}/ob^{2J}. These data show that the polymorphism (*i.e.*, mutation) shown in Figure 15A segregates in a genetic sense.

cDNa Cloning and Sequence Determination of *ob*

Using the labeled 2G7 PCR probe, a total of 50 mouse cDNA clones from a murine fat cell λgt11 cDNA library (Clonetech 5′-STRETCH cDNA from testicular fat pads of Swiss mice, #ML3005b), and thirty cross hybridizing human cDNA clones from a human fat cell λgt10 cDNA library (Clonetech 5′-STRETCH cDNA from abdomen #HL1108a) were isolated. Library screening was performed using the plaque lift procedure. The filters from the plaque lift were denatured using the autoclave method. The filters were hybridized in duplicate with the PCR labeled 2G7 probe (Rapid Hybe buffer, 65°C, overnight). After a 2-4 hour prehybridization, the filters were washed in 2x SSC, 2% SDS, twice for 30 minutes at 65°C and exposed to SRy Llim. Duplicate positives were plaque purified. Plaque purified phage were PCR amplified using commercially available vector primers, *e.g.*, λgt10 and λgt11. The resulting PCR products corresponded to the cDNA insert for each phage with a small amount of vector sequence at either end. The bands were gel purified and sequenced using the ABI automated sequencer and the vector primers to probe the DNA polymerase.

The raw sequencing data were then manually examined base by base to correct mishearing from the computer program. As the correct sequence became available, the downstream primers were synthesized and used to continue sequencing. Such experiments were repeated until each available cDNA clone was sequenced and synthesized into a contig. To date, ~3000 base pairs from the 5′ end of the mRNA has been compiled. One of the cDNA clones extended to the 5′ end of the mRNA since its sequence was identical to that of the 5′ RACE product of fat tissue RNA (data not shown).

The sequence data revealed that there is a 167 amino acid open reading frame (Figure 1). A Kozak translation initiation consensus sequence was present with an adenosine residue three bases upstream of the ATG. Two classes of cDNA were found differing by inclusion or exclusion of a single glutamine codon. This residue is found in a position immediately 3′ to the splice acceptor of the 2G7 exon. Since the CAG codon of glutamine includes a possible AG splice acceptor sequence, it appears that there is slippage at the splice acceptor site with an apparent 3 base pairs deletion in a subset of the cDNA, as shown below.

```
                gln  ser  val
          ag CAG TCG GTA  (with glutamine)          (SEQ ID NO:16)
             ↑
(splice acceptor site)

                ser  val
          ag CAG TCG GTA  (without glutamine)       (SEQ ID NO:17)
                ↑
          (splice acceptor site)
```

The "ag" in the sequences above corresponds to the assumed intron sequence upstream of the glutamine codon, and *AG* is the putative alternative splice site. This glutamine residue is located in a highly conserved region of the molecule and its importance for biological activity is as yet unknown.

A putative N-terminal signal sequence was detected, the signal cleavage site of which is predicted to be carboxy terminal to the alanine residue at amino acid position 21. This putative signal sequence was confirmed by application of a computer algorithm to the method of von Heijne (*Nucl. Acids Res.* **14**, 4683, 1986). Using this technique, the most probable signal sequence was identified in the polypeptide coding region corresponding to amino acids 1-23, having the sequence:

MCWRPLCRFLWLWSYLSYVQA ↑ VP (SEQ ID NO:10)

in which the arrow indicates the putative signal sequence cleavage site. The rest of the amino acid sequence was largely hydrophilic and did not have any notable

structural motifs or membrane spanning domains other than the N-terminal signal sequence. Specifically, we did not find consensus sequences for N-linked glycosylation or dibase amino acid sequences indicative of protein cleavage in the predicted processed protein (Sabatini and Adesnik, *The metabolic basis of inherited disease*, C.V. Scriver et al. eds., McGraw-Hill: New York, pp. 177-223). Data base search using Blast and Block programs did not identify any homologous sequence.

Human fat tissue RNA was analyzed on Northern blot, RNA species of similar size to the mouse *ob* gene was detected. Sequencing and analysis of cDNA clones revealed that human *ob* also encodes 167 amino acid polypeptide (Figures 2 and 3). Two classes of cDNA with or without three base pairs deletion were found in human as well (Figure 6). The mouse and human *ob* genes were highly homologous in the predicted coding region, but had only 30% homology in the available 3′ and 5′ untranslated regions. An N-terminal signal sequence was also present in the human *ob* polypeptide. Comparison of the human and mouse *ob* polypeptide sequences showed that the two molecules share an overall 84% identity at amino acid level (Figure 4). The N-termini of the mature proteins from both species share even higher homology, with only four conservative and three nonconservative amino acid substitutions among the N-terminal 100 amino acid residues.

Genomic DNA was isolated from mouse, rat, rabbit, vole, cat, cow, sheep, pig, human, chicken, eel, and drosophila, and restriction digested with EcoR1. The digests were electrophoresed on 1% agarose TBE gel. DNA was transferred to an imobilon membrane and probed with the PCR labeled 2G7 probe. The filter was hybridized at 65°C and washed with 2x SSC, 0.2% SDS at 65°C twice for twenty minutes each wash, *i.e.*, there were two buffer changes. These data indicate that *ob* is conserved among vertebrates (Figure 16). Note in this regard that there is a 2+ signal in eel DNA; eel is a fish.

In summary, available evidence suggests that body weight and adiposity are physiologically controlled. Seven years ago efforts began to identify two of the key components of this system: the *ob* and *db* genes. As shown in this example, the *ob* gene has now been identified as a fat specific gene that plays a key role in regulating body weight. The product of this gene, which is most probably a secreted hormone, will have important implications for the diagnosis and treatment of nutritional disorders in man and non-human animals.

EXAMPLE: Expression of *ob* In Bacteria

Both murine and human cDNAs encoding *ob* have been cloned into a pET-15b expression vector (Novagen). This vector contains a T7 promoter in conjunction with a lac operator, and expresses a fusion protein containing a histidine tag (His-Tag) and a thrombin cleavage site immediately upstream of the coding sequence insertion site (Figure 17) (SEQ ID No:11).

The mouse and human cDNAs were modified such that the alanine at the end of the signal sequence was turned into an NdeI site, as was a separate sequence in the 3′ region. Insertion of the NdeI site was accomplished using PCR with novel primers:

Mnde-5′ (murine five prime primer):
CTTATGTTCA TATGGTGCCG ATCCAGAAAG TC (SEQ ID NO:12)

Mnde-3′ (murine three prime primer):
TCCCTCTACA TATGTCTTGG GAGCCTGGTG GC (SEQ ID NO:13)

Hnde-5′ (human five prime primer):
TCTATGTCCA TATGGTGCCG ATCCAAAAAG TC (SEQ ID NO:14)

Hnde-3′ (human three prime primer):
TTCCTTCCCA TATGGTACTC CTTGCAGGAA GA (SEQ ID NO:15)

The primers contain a 6-base pair mismatch in the middle that introduces NdeI restriction sites at each end of the PCR fragment. Phage carrying either the mouse or human cDNA were PCR amplified using those primers. The PCR

product was digested with NdeI and gel purified on a 1% low melting point agarose gel. The gel purified bands were subcloned into the pET vector. The resulting plasmids were sequenced to ensure that mutations were not introduced during the PCR amplification step of cloning. Constructs for the human and murine cDNA that encodes and that lacks glutamine 49 have been prepared. In particular, pET 15b constructs containing either the human or the mouse *ob* coding sequence, minus signal sequence and fused to a Hig-Tag, have been made using a PCR cloning method. The constructs have been sequenced to ensure no sequence errors were introduced into the coding region of the *ob* gene during the PCR amplification step.

Two resultant plasmid constructs, pETM9 and pETH14, were selected to transform a bacterial expression host. Upon induction with 1 mM IPTG under optimal conditions, the transformed bacteria were able to produce 100-300 μg/ml of the *ob* fusion. The majority of the *ob* fusion protein was found in the inclusion body. After solubilization with 6M guanidine-HCl or urea, the fusion protein was purified through a His-binding (Ni-chelation) resin column. The conditions for column purification of the *ob* fusion protein (including binding, washing, and eluting) were established experimentally. The *ob* fusion protein binds to the resin at 5 mM imidazol/6M guanidine-HCl and stays bound at up to 20 mM imidazol/6M guanidine-HCl. The protein can be eluted form the resin at 60 mM imidazol/6M guanidine (Figure 18A,B). Both the purified human and mouse *ob* fusion proteins were further dialyzed in PBS to remove guanidine-HCl from the preparation and used to raise polyclonal antibodies.

In order to test the biological activity of the fusion protein products, the refolding conditions for the purified protein was tested and developed. This involves initial dialysis of the fusion protein in 1 M guanidine solution, followed by dilution with 0.4 M arginine solution. The His-Tag was removed from the fusion proteins before biological function assay. The tag removal was achieved by treating the fusion protein with thrombin from human placenta.

In addition, human and mouse *ob* gene coding sequence minus the signal sequence is being inserted into a pET 12c vector using PCR cloning method. These constructs can direct the synthesized *ob* fusion proteins into the periplasmic space of the bacterial host cell. The *ob* fusion protein recovered from the periplasmic space may only need a simple gel filtration to be purified from other host proteins and will not be denatured during such process.

EXAMPLE: Preparation of Antibodies to the ob Polypeptide

In addition to use of the recombinant protein to generate polyclonal antibodies, a set of four peptide sequences from the deduced murine ob sequence were identified using immunogenicity plot software (GCG Package). The four carboxyl terminal peptide fragments are:

(SEQ ID NO:18):
Val-Pro-Ile-Gln-Lys-Val-Gln-Asp-Asp-Thr-Lys-Thr-Leu-Ile-Lys-Thr

(SEQ ID NO:19):
Leu-His-Pro-Ile-Leu-Ser-Leu-Ser-Lys-Met-Asp-Gln-Thr-Leu-Ala

(SEQ ID NO:20):
Ser-Lys-Ser-Cys-Ser-Leu-Pro-Gln-Thr-Ser-Gly-Leu-Gln-Lys-Pro-Glu-Ser-Leu-Asp

(SEQ ID NO:21):
Ser-Arg-Leu-Gln-Gly-Ser-Leu-Gln-Asp-Ile-Leu-Gln-Gln-Leu-Asp-Val-Ser-Pro-Glu-Cys

These peptides were conjugated to KLH, and the peptide-KLH conjugates were used to immunize rabbits using standard techniques. Polyclonal antisera specific for each peptide is recovered from the rabbits.

EXAMPLE: *In Vitro* Translocation of an ob Polypeptide

In order to confirm the presence of a functional signal sequence, a human cDNA that included the entire open reading frame was subcloned into the pGEM vector.

Only the human cDNA was used in this experiment because suitable mouse subclones were not recovered. Positive strand human *ob* RNA was transcribed using sp6 polymerase and used in an *in vitro* translation reaction with and without canine pancreatic microsomal membranes. The primary translation product migrated with an apparent molecular weight of ~18 kD, which is consistent with that predicted by the cDNA sequence. Inclusion of the microsomal membranes in the reaction inhibited the overall efficiency of translation ~5 fold. Nevertheless, approximately 50-70% of the *ob* primary translation product was truncated by approximately 2 kD in the presence of the membrane preparation, suggesting that the signal sequence is functional (Figure 19A). The size of the primary translation product of interleukin-1α RNA, which does not encode a signal sequence, was unchanged when microsomal membranes were included in the reaction. In order to confirm that translocation of the *ob* protein had taken place, the *in vitro* translation products were treated with Proteinase-K. Protease treatment resulted in the complete proteolysis of the 18 kD primary translation product while the 16 kD processed form was unaffected by the enzyme treatment, indicating that it had translocated into the lumen of the microsomes (Figure 19B). These data are compatible with the hypothesis that *ob* is a secreted molecule.

After signal sequence cleavage, two cysteine residues would remain within the predicted protein raising the possibility that the molecule contains a disulfide bond characteristic of other secreted polypeptides (Shen and Rutter, 1984, Science 224:168-171).

EXAMPLE: Characterization of the *ob* Gene

To establish the relationship between obesity and genetic alterations in the *ob* gene in humans, the sequence of the human *ob* gene was determined (Figure 20A) (SEQ ID NO:). Specific primers from the human coding sequence were used to screen a human P1 library. Three different P1 clones were obtained, grown up, and PCR amplified using primers flanking the splicing site between the first and second

coding exon. The entire intron region, around 2 kB, was amplified and partially sequenced (see Figure 20A; and as indicated in SEQ ID NO:22).

The gene structure of both the murine and human genes was characterized using PCR assays and other standard techniques. The mouse *ob* gene was found to consist of 3 exons, the second and third of which account for the coding sequence (Figure 20B). The coding region of the human *ob* gene shares the same structure; however, the human gene lacks a 5′ exon and intron (Figure 20C).

Two sets of primers generated from the intronic sequences of the human gene have been prepared (Figure 20A). The sequences of the primers follows (F and R refer to forward and reverse, respectively):

HOB 1gF	5′-CCCAAGAAGCCCATCCTG-3′ (SEQ ID NO:26)
HOB 1gR	5′-GACTATCTGGGTCCAGTGCC-3′ (SEQ ID NO:27)
HOB 2gF	5′-CCACATGCTGAGCACTTGTT-3′ (SEQ ID NO:28)
HOB 2gR	5′-CTTCAATCCTGGAGATACCTGG-3′ (SEQ ID NO:29)

DNA samples have been obtained from various sources, and these sets of primers are being used to amplify human genomic DNA from severely obese people. The PCR products were run on a low melting point agarose gel, and the bands were cut out and digested with agarase. The sequences were obtained using the ABI 373A DNA sequencer and Taq dideoxy terminator kit (abi, Perkin-Elmer). One point mutation in an *ob* gene from a patient sample has been detected to date. This mutation is on the first exon and does not change the amino acid sequence. Preliminary data indicate that an insertion sequence may be present in the first exon of another patient.

A different automated sequencing method with Sequenase instead of Taq DNA polymerase may be employed to yield more easily readable sequences for mutation detection.

EXAMPLE: Expression of *ob* in Yeast

Following the positional cloning of *ob*, it became important to uncover the physiological mechanism by which the *ob* protein reduces food intake and body weight. The first step in this direction was to recombinantly produce a functional protein using an expression system. In addition to the successful bacterial expression system, a yeast expression system was also selected. Yeast expression has several attractive features for expressing *ob*. The most important is that biologically active eukaryotic proteins are more likely to be produced. The ob polypeptide is secreted by mammalian cells. Protein secretion is very similar for all eukaryotes, which means that the yeast secretory apparatus is much more similar to the mammalian secretory pathway than bacterial secretory pathways would be. In particular, protein modifications of *ob* seen in mammalian cells would likely also be seen in the expression through the yeast secretory system. In addition, protein folding is carried out in passage through the secretory apparatus and thus delivering ob through the yeast secretory apparatus is likely to give a properly folded protein with native biological activity. This is significant for *ob* because the two cystein residues may form a disulfide bridge. In contrast to secretory pathways, the reducing environment of the cell cytoplasm prevents formation of disulfide bridges, and therefore it is essential that ob pass through the secretory pathway in order for this disulfide bond to form *in vivo*. Other advantages have to do with the ease and quickness of manipulating yeast, the availability of vectors and strains, and the vast experience in yeast recombinant technology.

A *Pichia pastoris* expression system was chosen for four reasons: (1) it has higher levels of heterologous protein expression than other yeast systems such as *S. cerevisiae*; (2) protein glycosylation is more similar to the mammalian system in *P. pastoris* than *S. cerevisiae* (although glycosylation sites were not detected in ob using a computer search, there still remained the possibility of glycosylation at unrecognized sites); (3) *P. pastoris* secretes very few proteins natively, and thus it

is generally straightforward to purify the expressed foreign protein; and (4) the vectors and yeast strains are commercially available (from Invitrogen). Two strategies for generating yeast expression vectors are shown in Figures 21 and 22.

The vector chosen was pPIC.9. This vector contains a cloning site just downstream of the alpha-mating factor prepro coding sequence which directs the protein encoded by the gene cloned into the cloning site to be secreted by the secretory pathway. The other important feature of the vector is a HIS4 gene that allows selection for uptake of the vector using a yeast auxotrophic strain grown on histidine-deficient media following transformation of the yeast with the vector. The cloning strategy was the following: PCR amplify *ob* cDNA using a 5′ primer that contained at its 3′ end sequence complementary to the sequence of *ob* just following the predicted leader peptide cleavage site, and at its most 5′ end a sequence complementary to the 3′ end of the alpha-mating factor sequence of the vector. The 5′ primer also contains an XhoI site. The 3′ primer was designed to have at its 3′ end a sequence complementary to the last few amino acids of *ob* and an EcoRI site at its 5′ end. Following PCR amplification, the PCR product was digested with XhoI and EcoRI and cloned into similarly digested pPIC.9. Following the cloning of both the mouse and human *ob* cDNAs, each with and without the glutamine at codon 49, individual clones were isolated for all four individual constructs and sequenced to verify that the constructs were cloned in the correct orientation and frame and contained no mutations from the PCR amplification step. Following identification of clones with the correct sequence, these were transformed into *P. pastoris* strain GS115, a histidine auxotroph.

For the two mouse *ob* constructs, transformed yeast clones were screened for protein expression. As evidence that the transformed yeast contain *ob*, a DNA dot-blot assay and a colony hybridization assay were done which both showed *ob* sequence within the transformed yeast but not within the untransformed yeast. Furthermore, the transformed yeast now secreted a 16 kDa protein into the culture media whereas the untransformed yeast does not secrete a protein of this size

(Figure 23A). This is the predicted size of *ob*. Individual clones for both mouse constructs have been identified that are high expressors for *ob*, and currently a purification strategy is being developed to purify *ob* to homogeneity. One strategy has been to purify *ob* on a cation exchange column (Figure 23B); preliminary data suggest that a strong cation exchanger may be useful. However, after cation exchange chromatography, the putative ob product is lost. This indicates the presence of a protease in the sample.

One strategy to overcome this problem is to prepare *ob*-His tag fusions for expression in yeast (Figure 22). Further evaluation has demonstrated that *ob* without a His tag associates tightly with a Ni-chelation column. Purification of the ob polypeptide by Ni-chelation, followed by gel filtration, yielded a product of sufficient purity for mass spectral analysis. Mass spec confirms the molecular weight of the expressed protein is identical to the expected molecular weight, which strongly confirms that *ob* has been successfully expressed in *Pichia*.

However, the Ni-chelation/gel filtration purification protocol does not yield a ob polypeptide in sufficiently pure form. Additional small molecules are present. It does appear that the proteolytic activity elutes from the Ni-chelation column in the void volume. Accordingly, a three step purification process is planned: Ni-chelation, followed by cation exchange (which eliminates the small molecule contaminants), followed by gel filtration.

Estimating expression level by Coomassie blue staining of SDS-PAGE gels reveals approximately 10 mg/L when yeast are grown in shaker flasks. These levels are expected to increase in fermentation vessels, and we are about to initiate fermentation with the hopes of obtaining larger quantities of protein. Regarding the human *ob* constructs, transformed yeast clones containing high copy numbers of the *ob* gene have been identified, and these are expected to express ob protein. As antibodies are developed, these will be used to confirm the identity of the secreted 16 kDa protein.

Once purified the expressed protein is characterized by several methods. Physical characterization includes light-scattering to determine homogeneity of structure and is used as a measure of proper folding. Circular dichroism is used to roughly determine the structural geometry of the protein. Importantly, bioactivity of the purified protein is assayed by administering the protein to both elan and obese rodents via an osmotic pump (*e.g.*, an ALZET osmotic pump from Alza Corporation, Palo Alto, CA) over at least a two-week period and observing effects on feeding behavior and body weight.

The following is a list of references related to the above disclosure and particularly to the experimental procedures and discussions.

Bahary, N.; G. Zorich; J.D. Pachter; R. L. Leibel; and J.M. Friedman. 1991. Molecular genetic linkage maps of mouse chromosomes 4 and 6. Genomics 11:33-47.

Bahary, N.; D. McGraw; R.L. Leibel; and J.M. Friedman. 1991. Chromosomal microdissection of midmouse chromosome 4: Mapping of microclones relative to the mouse *db* gene. Submitted.

Bahary, N.; J. Pachter; R. Felman; R.L. Leibel; K.A. Albright; S. Cram; and J.M. Friedman. 1991. Molecular mapping of mouse chromosomes 4 and 6: Use of a flow-sorted Robertsonian chromosome. Submitted.

Blank, R.; J. Eppig; F.T. Fiedorek; W.N. Frankel; J. M. Friedman; K. Huppi; I. Jackson; and B. Mock. 1991. Mouse chromosome 4. Mammalian Genome 1(suppl): s51-s78.

Bogardus, C.; Ravussin, E.; Abbot, W.; Zasakzku, J.K.; Young, A.; Knowler, W.C.; Friedman, J.M.; R. L. Leibel; N. Bahary; D. A. Siegel; and G. Truett, G.

1991. Genetic analysis of complex disorders: Molecular mapping of obesity genes in mice and humans. Annals of the New York Academy of Sciences 630:100-115.

Friedman, J.M.; R. L. Leibel; and N. Bahary. 1991. Molecular mapping of obesity genes. Mammalian Genome 1:130-144.

Friedman, J.M.; R.L. Leibel; N. Bahary; and G. Zorich. 1991. Molecular mapping of the mouse *ob* mutation. Genomics, (in press).

Harris, M.I. (1991). Diabetes Care *14* (suppl. 3), 639-648.

Harris, M.I.; Hadden, W.C.; Knowler, W.C.; and Bennett, P.H.(1987). Diabetes *36*, 523-534.

Harris, R.B.S. (1990). FASEB J. *4*, 3310-3318.

Jacobowitz, R., and Moll, P.O. (1986). N. Engl. J. Med. *315*, 96-100

Kessey, R.E. (1980). In Obesity, A. Stunkard, eds. (Philadelphia: W.B. Sauders Co.), pp. 144-166.

Kessey, R.E., and Pawley, T.L. (1986). Annu. Rev. Psychol. *37*, 109-133.22

Leibel, R.L., N. Bahary and J.M. Friedman. 1990. Genetic variation and nutrition in obesity: Approaches to the molecular genetics of obesity. In Genetic variation and Nutrition (Simopoulos, A.P. and Childs, B., eds.), S. Karger, Basel, pp. 90-101.

Siegel, D.; N.G. Irving; J.M. Friedman; and B. J. Wainwright. 1991. Localization of the cystic fibrosis transmembrane conductance regulator to mouse chromosome 6. Cytogenetics Cell Genetics, submitted.

Truett, G.E.; N. Bahary; J. M. Friedman; and R. L. Leibel. 1991. The rat obesity fatty (fa) maps to chromosome 5:Evidence for homology with the mouse gene diabetes (db). Proc. Natl. Acad. Sci. USA 88:7806-7809.

THE REMAINDER OF THIS PAGE IS INTENTIONALLY LEFT BLANK

EXAMPLE: Weight Reducing Effects of the Ob Polypeptide (Leptin)

(Various references are cited by author, year, and # in this Example, which citations correlate with the list of references found at the end of this Example.)

The gene product of the mouse *ob* locus plays an important role in regulating body weight. We establish that the *ob* protein circulates in mouse, rat and human plasma. The circulating form in all three species has an identical molecular weight by SDS-PAGE to the deduced polypeptide sequence without the signal sequence, suggesting that *in vivo* the protein is not processed after cleavage of the signal sequence. The *ob* protein is absent in plasma from C57/Bl6J ob/ob mice and is present at ten-fold higher concentrations in plasma of db/db mice and twenty-fold higher levels in plasma of fa/fa rats relative to controls. These obese animal mutants have been suggested to be resistant to its effects. There were seven fold differences in plasma levels of the *ob* protein within a group of six lean human subjects. Daily injections of the recombinant mouse *ob* protein dramatically reduce body mass in ob/ob mice, have significant effects on body weight of wild type mice but have no effect on db/db mice. These data show that the gene product of the *ob* locus serves an endocrine function to regulate body weight. We propose that the protein encoded by the *ob* gene be named Leptin derived from the Greek root *leptós* meaning thin.

Materials and Methods

Rabbits were immunized with recombinant protein in Freunds adjuvant. (HRP, Inc) Immunopurified anti-mouse *ob* antibodies were prepared by passage of antiserum over a sepharose 4B column conjugated to the recombinant protein as described[Harlow, 1988 #444]. Immunoprecipitation of mouse plasma was carried out as follows. 0.5 ml of plasma from mouse, rat and human containing approximately 2.5 mM EDTA was pre-cleared with unconjugated sepharose-4B at room temperature with rocking for 2 hours. The sepharose was removed by spinning and 50 µl added of a 50% slurry of antibody-conjugated sepharose containing affinity purified antibody at a concentration of 1 mg/ml of packed sepharose. 0.5 ml of 2x RIPA buffer was added to give final binding conditions as follows: 50 mM Tris-HCl, pH7.5, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate and 0.025% sodium azide. The reaction was carried out overnight at 4°C with rocking. The antibody-conjugated sepharose was washed 8 times with RIPA buffer and rinsed three times with PBS and run on 15% SDS-PAGE. The proteins were

transferred to nitrocellulose and Western blotted with a biotinylated immunopurified antibody against the recombinant protein The secondary used was HRP-streptavidin and ECL was used for detection.

To quantitate the amount of *ob* in mouse serum, increasing amounts of the refolded recombinant mouse *ob* protein (0.01, 0.1, 0.5, 2.0, 15.0 ng) was added to 100 λ of C57BL/6J ob/ob plasma and incubated at 4°C for 3 hours with the protein A sepharose conjugated antibody. After extensive washing with buffer A (10 mM Na Phosphate buffer, pH 7.4; 100 mM NaCl; 1% Triton X-100, 5 mM EDTA, 1 mM PMSF) samples were resuspended in sample buffer and loaded on a 15% SDS-PAGE and transferred to a nitrocellulose membrane. Western blotting was performed using an immunopurified biotinylated anti-amino terminus antibody as a primary antibody and HRP-Streptavidin as a secondary antibody followed by ECL detection.

Cytoplasmic extracts were prepared by homogenizing adipose tissue in NDS buffer by polytron and dounce homogenization and removal of nuclei by centrifuging at 700 g. [10 mM Tris, pH 7.5, 10 mM NaCl, 60 mM ICCl, 0.15 mM spermine, 0.5 mM spermidine, 14 mM β-Mercaptoethanol, 0.5 m EGTA, 2 mM EDTA, 0.5% NP-40]

Immunoprecipitations were performed as above except that immunopurified anti-human *ob* antibodies were used. For the ELISA, 100 μl of a 1 μg/ml solution of immunopurified anti-human *ob* antibody was dissolved in a borate buffered PBS solution and applied overnight to microtiter (Corning cat. #2595) plates at 4° C. The plates were then washed 4 times with borate saline solution containing 0.05% Tween 20 and excess liquid was removed. Plates were blocked by incubation at RT for 2 hours with 240 μl per well of borate saline buffer containing 0.3% gelatin and then washed and dried. Either known amounts of a refolded human ob protein or plasma samples in 100 μl volume were incubated in individual wells overnight at 4° C. After washing, the plates were incubated with 100 μl of a biotinylated immunopurified anti-human antibody (0.1 mg/ml in a gelatine borate buffered solution) for 4 hours at room temperature. After washing Horse Radish Peroxidase-Streptavidin was added to the plates (0.1 μg/ml in borate buffer, 0.3% gelatin). HRP substrate solution (ABTS, 0.3 mg/ml and H_2O_2, 0.01% in citric acid) was then used for detection and the OD at 414 nM was read to quantitate the antibody binding.

The mouse and human *ob* coding sequence were PCR amplified from plasmids containing *ob* cDNA sequences and subcloned into the pPIC.9 plasmid (Invitrogen). The human 5' primer was: 5' GTATCTCTCGAGAAAAGAGTGCCCATCCAAAAAGTCCAAG 3' and the 3' primer was 5' GCGCGAATTCTCAGCACCCAGGGCTGAGGTC 3'. For mouse the 5' primer was: 5' GTATCTCTCGAGAAAAGAGTGCCTATCCAGAAAGTCCAGG 3' and the 3' primer was 5' GCGCGAATTCTCAGCATTCAGGGCTAACATC 3'. The 5' primer for both mouse and human

contains an XhoI site at the 5' end and coding sequences for the last 4 aa of the alpha-mating factor signal sequence present in the vector pPIC.9. This vector directs secretion of heterologously expressed genes from the cell into the culture media. The 5' PCR primer also includes the first 19 n.t.'s of the *ob* open reading frame after the signal sequence cleavage site before the alanine at amino acid position 21. The 3' primer contains an EcoRI site at it's 5' end which is immediately followed by sequences complementary to the putative *ob* stop codon. The PCR conditions were as follows: denaturing for 1' at 94, annealing for 1' at 55 and extension for 2.5' at 72. Low-cycle PCR (15 cycles) and the proof-reading polymerase PFU (Stratagene) were used to limit the number of PCR-generated mutations. The PCR products were digested with XhoI and EcoRI and cloned into similarly digested vector pPIC.9. All constructs were sequenced on both strands to ensure the absence of any PCR-generated mutations. Clones were transformed into Pichia pastoris (His-) by the spheroplast method and selected on histidine deficient media. Approximately 200 clones of mouse and human were screened for high-copy number integration by a colony hybridization assay and the high copy number clones were then assayed for *ob* expression initially by coomassie staining showing the presence of a novel 16 kd protein present in the culture media of transformed yeast. The 16 kd band was confirmed to be *ob* using antibodies raised against the bacterially expressed protein. The recombinant proteins were purified by a two-step purification method described below. Mass spectrometry and cyanogen bromide treatment were performed as described[Beavis, 1990 #804].

The entire ob coding sequence of the mouse and human ob genes C-terminal to the signal sequence were subcloned into the Pet15b expression vector (Novagen) and overexpressed in *Escherichia coli* [BL21(DE3)plYsS] using the T7 RNA polymerase system[Studier, 1990 #803]. Cells grown at 30°C to an absorbency of 0.7 at 595 nM and induced with 0.5 mM isopropyl-β-D-thiogalcto-pyranoside overnight were collected by low-speed centrifugation. Lysis was performed by three cycles of freeze thaw and DNA digestion was done with DNaseI. Membrane extraction was performed by sonication and detergent solubilization, and the final inclusion body pellet was dissolved in 6M guanidine-HCl, 20mM HEPES, pH8.4. Recombinant ob proteins were purified under denaturing conditions by IMAC using a Ni-ion affinity column and washing with increasing amounts of imidazole. Purified denatured *ob* protein was then stored in 6 M guanidine-HCl, 10 mM sodium acetate (NaAc), pH5, and reduced using 1 mM DTT at room temperature for 1 hour. Denaturation was performed by diluting the reduced protein into 20% glycerol, 5 mM $CaCl_2$, 5 mM NaAc, pH5, through mixing and incubation at room temperature for 8-12 hours. After denaturation the pH was adjusted to 8.4 by addition of Tris to 10 mM, and the hexahistidine tag was removed by thrombin cleavage. Cleaved, renatured protein was repurified by IMAC to separate product from thrombin and uncleaved fusion protein.

Cleaved, renatured protein elutes from the Ni-ion affinity column at 40 mM imidazole, whereas thrombin is not retained and uncleaved fusion protein elutes at 0.2 mM imidazole. Product was then concentrated, treated with 100 mM EDTA and 10 mM potassium ferricyanide and further purified by gel filtration using Pharmacia superdex 75 16/60 column.

An Ellman's assay was conducted as described[Ellman, 1959 #798]. Ellman's reagent was prepared by dissolving 39.6 mg 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB) in 10 ml 0.05 M phosphate, pH 8. A calibration curve was constructed in the concentration range of 10-120 mM free sulfhydryl (using a 1 mM stock solution of reduced DTT) at 412 nm. Each assay was performed using 0.02 ml Ellman's reagent and a total reaction mixture of 0.5 ml. The measured extinction coefficient was 12974 $M^{-1}cm^{-1}$ for free sulfhydryl group (correlation coefficient 0.99987), which is within 5 % of the previously reported value of 13600 $M^{-1}cm^{-1}$.

50 µl of 2 µg/ml pure gel filtered protein, corresponding to a possible free sulfhydryl concentration of about 24 µM in the final reaction mixture, was subjected to Ellman's assay. The resulting solution gave A_{412} of about 0.02, suggesting that two cysteine residues in the protein are in oxidized state to form cystine or that their free sulfhydryl groups are completely buried within the inaccessible core of the folded protein. Identical results were obtained by conducting the same assay on unfolded protein in the presence of 6 M guanidine-HCl.

Mice were individually caged in a pathogen-free environment and acclimated to a diet containing 35% (w/w) Laboratory Rodent Diet 5001 (PMP Feeds, Inc.), 5.9% (w/w) tapioca pudding mix (General Foods) and 59.1% water which has an energy content of 1.30 kcal/gm. The diet was sterilized by autoclave and packed into 60 mm plastic dishes which were fixed to the tops of 100 mm petri dishes. Tapioca gives the diet a pasty texture making it difficult for the animal to spread the food in the cage. The 100 mm lid recovers the small amount of food spilled by the animal. A fresh dish of food was placed into the cage each morning and the previous day's dish was removed and weighed. The difference in weight provided a measure of daily food consumption. Effects of recombinant protein on food intake and body weight were measured in three strains of mice: C57Bl/6J ob/ob, C57 Bl/Ks db/db and CBA/J +/+, purchased from the Jackson Laboratory. Thirty mice from each strain were divided into groups of 10. One group from each strain received daily intraperitoneal (i.p.) injections of the refolded bacterial *ob* protein at a dose of 5 µg/g/day in 300 µl of PBS A second group received i.p. injections of the same volume of PBS. These control mice received injections of the PBS dialysate of the recombinant protein. The PBS was cleared of endotoxin using an Acticlean ETOX column. A third group of animals did not receive injections. Food intake was recorded daily and body weight measurements were recorded regularly over a 3.5 week interval. For the pair feeding experiment, the food intake of a separate group of *ob* mice was matched on a daily basis to that consumed by the *ob* mice receiving protein.

RESULTS

The ob Protein Circulates in Mouse, Rat and Human Plasma

Recombinant mouse and human *ob* protein was prepared using the PET 15b bacterial expression vector (Novagen) and by cloning into Pichia pastoris, a yeast expression system that secretes recombinant proteins directly into the culture media. The *ob* protein expressed in yeast includes the 146 amino acids carboxy terminal to the signal sequence. Rabbits were immunized with the bacterial proteins (HRP, Inc.). Antibodies were immunopurified (Research Genetics) and used for immunoprecipitations and Western blots of protein from plasma and adipose tissue.

The *ob* protein from mouse plasma migrates with an apparent molecular weight of 16 kD by SDS-PAGE. The electrophoretic mobility is identical to the recombinant *ob* protein secreted by yeast after signal sequence removal. (Fig. 2A) The protein was not detected in plasma from C57BL/6J ob/ob mice that have a nonsense mutation at codon 105. Several different antisera failed to identify the truncated 105 residue polypeptide chain predicted by the cDNA sequence.

A ten-fold increase in the level of circulating protein was observed in db/db mice relative to a control animal (Fig. 2A). Immunoprecipitation of plasma from wild type and *fa/fa* rats revealed a twenty-fold increase in the level of the *ob* protein in the mutant rat compared to wild type. (Fig. 2B) The *db* mutation results in an obese phenotype identical to that seen in *ob* mice[Bahary, 1990 #31]. fatty rats are obese as a result of a recessive mutation in a gene homologous to *db* [Truett, 1991 #409]. In order to quantitate the level of *ob* in mouse plasma, increasing amounts of recombinant protein were added to *ob* serum and immunoprecipitated. (Fig. 2C) A linear increase of the signal intensity on Western blots was seen with increasing amounts of recombinant protein. Comparison of the signal intensity of the native protein in mouse plasma to the standards indicated that the circulating level of the *ob* protein in wild type mice is ~20 ng/ml. These data demonstrate that the immunoprecipitations and Western blots were performed under conditions of

antibody excess. Increased levels of the *ob* protein were also seen in protein extracts of adipose tissue from db/db mice relative to controls. (Fig.24D) As expected for a secreted protein, the protein from the adipose tissue fractionated with the crude membrane fraction (data not shown).

Plasma samples from six lean human subjects with a Body Mass Index less than 25 (BMI=weight/length^2) were immunoprecipitated using immunopurified antibodies to the human protein. The immunoprecipitated material migrated with an electrophrotic mobility identical to that seen for the 146 amino acid human protein expressed in yeast. The intensity of the signals varied significantly among the six samples. (Fig 25A) Densitometry of the autoradiograph revealed an ~ five-fold difference in the levels in individuals HP1 and HP6 with intermediate levels in the other subjects. An enzyme linked immunoassay (ELISA) was developed using the immunopurified antibody and the refolded bacterial protein as a standard (see below). The resulting standard curve is shown in Fig.25B. Using this assay, the plasma levels of the *ob* protein in the six human plasma samples varied between 2-15 ng/ml. (Fig.25C) The level of the *ob* protein in plasma from HP 6 was outside of the linear range of the immunoassay and is $\geq$ or 15 ng/ml. These quantitative differences correlated with those seen on Western blots.

Structural Features of the *ob* Protein

As the *ob* protein has two cysteine residues, it could form either intra- or intermolecular disulphide bonds under oxidizing conditions in vivo. Western blots were repeated with and without the addition of reducing agents to the sample buffer. Under both conditions, the *ob* protein in human serum migrated as a monomer (data not shown). Under nonreducing conditions, protein immunoprecipitated from *db* mouse serum was detected at positions consistent with that of both a monomer of 16kD and a dimer of ~32kD. (Fig.26A) The higher molecular weight moiety disappeared under reducing conditions suggesting that a fraction of mouse *ob* circulates as a higher molecular weight species via formation of an intermolecular disulphide bond. Approximately 80% of mouse *ob* circulates as the ~16 kD protein and 20% as the ~32 kD form.

The same molecular forms are seen when the mouse and human proteins are expressed in Pichia pastoris [Abrams, 1992 #708]. In these studies, the DNA sequence corresponding to the 146 amino acid mature *ob*

protein was cloned downstream of the yeast alpha mating factor signal sequence in the pPIC.9 vector (Invitrogen). *ob* protein was purified from the yeast media of strains expressing the mouse and human proteins and electrophoresed under reducing and nonreducing conditions. (Fig. 26A) The mouse protein was expressed in yeast mainly as a dimer under nonreducing conditions, and only as a monomer in the presence of reducing agents. The recombinant human protein migrated to the position of a monomer under both conditions (data not shown).

The purified human protein expressed in Pichia had a molecular mass of 16,024±3 Da as determined by mass spectrometry[Beavis, 1990 #804]. This value is in agreement with the mass calculated from the amino acid sequence of the protein containing a single intramolecular disulfide bridge (16,024 Da). Matrix-assisted laser desorption mass spectometric analysis of cyanogen bromide cleavage products of the protein indicates that cysteines 117 and 167 are linked through an intramolecular disulphide bond. (Fig. 26B) Cyanogen bromide cleaves carboxyterminal to methionine residues.

Preparation and Characterization of Bioactive Recombinant Protein

Mouse ob protein was expressed in *E. coli.* from a PET 15b plasmid as an insoluble fusion protein, with a twenty residue, N-terminal hexahistidine tag containing a thrombin cleavage site. Bacterial inclusion bodies were solubilized using guanidine-HCl and purified under denaturing conditions using immobilized metal ion affinity chromatography (IMAC). (Fig. 27) Purified, denatured fusion protein was reduced, diluted and permitted to refold in aqueous solution at room temperature. Following thrombin cleavage, renatured mouse *ob* protein containing four additional N-terminal residues (Gly-Ser-His-Met) was repurified by IMAC to >98% homogeneity, as judged by SDS-PAGE and mass spectrometry. Matrix-assisted laser desorption mass spectrometry gave a measured mass of 16,414±3 Da (predicted mass = 16,415 Da). Both reducing and non-reducing SDS-PAGE gels demonstrated a single molecular species with apparent and molecular weight of 16 kD (data not shown).

Dynamic light scattering using a DP801 Molecular Size Detector (Protein Solutions, Inc.) demonstrated that the renatured mouse ob protein was largely monomeric, with some higher-order aggregates. The protein was treated with

EDTA and chemically oxidized. Higher molecular weight species were then removed by gel filtration.

Further dynamic light scattering confirmed that the purified, renatured recombinant mouse ob protein was monodispersed. Following dialysis against phosphate buffered saline (PBS), bacterial endotoxin was removed using an Acticlean ETOX column (Sterogene Bioseparations, Inc.). The final yield of protein was 45 mg/l.

Ellman's assay was performed on the purified, renatured recombinant mouse *ob* protein to assess its oxidation state[Ellman, 1959 #798]. Both renatured protein and protein unfolded by 6M guanidine-HCl demonstrated <0.5% free sulfhydryl content, demonstrating that the monomeric product contains an intramolecular disulphide bond. This was confirmed by mass spectrometry of the cyanogen bromide cleavage products of the refolded bacterial protein (data not shown).

Bioactivity of the ob Protein

The purified, renatured recombinant mouse *ob* protein was administered as a daily intraperitoneal injection of 5 mg/kg/day to groups of 10 C57Bl/6J ob/ob (age, 16 weeks), C57Bl/Ks db/db (age, 12 weeks) and CBA/J +/+ (age, 8 weeks) mice. An equal number of animals received PBS as a daily injection. The PBS used for the control injections was derived from the dialysate after equilibration of the protein. Ten additional animals from the three mouse strains did not receive injections. The food intake of individual animals was monitored daily and the weights of the animals were recorded at three or four day intervals. The cumulative results for food intake and body weight from each of the 9 groups of mice are shown in Fig.28A and the statistical significance of the data are shown in Table 1. The food intake of the C57Bl6J ob/ob mice injected with protein was significantly decreased after the first injection and continued to decrease until the fifth day when it stabilized at a level equal to ~40% of the intake of the animals receiving injections of PBS ($p < .001$). The sham injected *ob* mice did not lose weight over the three week study period. The C57Bl/6J ob/ob mice receiving protein lost approximately 10% of their body weight after 5 days ($p < .001$). These animals continued to lose weight over the three week treatment at which point the weight of the *ob* animals receiving protein had decreased to an average of 60% of their initial body weight (p <

.0001). A separate group of *ob* mice were pair fed to the *ob* mice receiving protein. The data in Fig. 2B show that the pair fed mice lost significantly less weight than the animals receiving the recombinant protein. ($p < .02$). A photograph of two mice receiving injections of either protein or vehicle shows the gross difference in appearance resulting from the protein treatment (Fig. 2C). In order to further ascertain the effects of the protein, autopsies of two mice in each of the groups were performed. Gross inspection of the *ob* mice receiving protein revealed a dramatic decrease in body fat as well as the size of the liver. The liver weights of the *db* and wild type mice were unchanged with treatment. The livers from the *ob* mice receiving the injections of PBS weighed 5.04 and 5.02 grams vs. 2.23 and 2.03 grams in the animals receiving the recombinant protein. In contrast to the pale fatty liver characteristic of *ob* mice, the liver from the *ob* mice receiving protein acquired the darker color characteristic of normal liver (Fig. 2D). Histologic sections of the liver indicated that the untreated animals had a fatty liver that was markedly improved by protein treatment (data not shown).

In contrast to the *ob* mice, there were no significant differences in body weight or food intake in the C57BL/Ks db/db mice receiving protein relative to the control group receiving vehicle. (Fig. 2A, Table 1) All three groups of db/db mice lost between 2-5 grams during the treatment period. The average blood glucose of the *db* mice was measured using a glucometer, and was $\geq$ 500 mg/dl in all of the mice indicating that these animals had developed diabetes secondary to obesity. The injections of *db* mice were terminated after two weeks.

In wild type mice there was a small but significant decrease in body weight following administration of the recombinant *ob* protein. (Fig. 2A, Table 1). After five days of protein injection, the treated mice lost an average of .5 grams while control mice gained .4 grams ($p < .02$). At two subsequent time points the animals receiving protein weighed significantly less than the mice receiving daily injections of PBS. The significance of the weight change was reduced at the later time points. In the animals that lost weight, the food intake was not significantly different from control animals. The injections of PBS had a small but significant effect on food intake and body weight in *ob*, *db* and wild type mice as compared to mice not receiving injections ($p < .05$).

Table 1.

Animal Group	Treatment Group	WEIGHT CHANGE Days	n	Mean	Std. Error	p
ob/ob	protein	1-5	10	-6.38000000	0.47628190	<0.001
	vehicle		9	-0.14444444	0.24444444	
	protein	1-12	10	-14.45000000	0.70793126	<0.001
	vehicle		9	0.98888889	0.38058597	
	protein	1-27	6	-24.28333333	0.69924563	<0.0001
	vehicle		5	4.30000000	0.79874902	
db/db	protein	1-5	10	-1.47000000	0.36939891	0.240
	vehicle		10	-2.00000000	0.23142073	
	protein	1-12	10	-3.75000000	0.77348418	0.610
	vehicle		10	-4.19000000	0.34655447	
CBA/J	protein	1-5	10	-0.48000000	0.17876117	0.006
	vehicle		10	0.38000000	0.21489015	
	protein	1-12	10	-0.12000000	0.45748103	0.015
	vehicle		10	1.20000000	0.18378732	
	protein	1-27	5	1.98000000	0.48723711	>0.651
	vehicle		6	2.23333333	0.20763215	

The effects of protein vs PBS injections on body weight are tabulated for C57Bl/6J ob/ob, C57Bl/Ks db/db and wild type mice. Shown are the average values, standard errors and statistical significance at each of the time points listed. Four animals in the CBA/J and *ob* groups and all of the *db* mice were sacrificed after two weeks of injections. Statistical significance was determined using a two sample of test.

DISCUSSION

An endocrine function for the protein product of the *ob* locus was first suggested by Coleman, who showed that the body weight of ob/ob mice was reduced after parabiotic union to normal or *db* mice[Coleman, 1978 #42]. Our results support this hypothesis by showing that *ob* protein circulates in the bloodstream and that injections of recombinant protein reduce body weight. The molecular weight of the gene product encoded by the *ob* gene is ~16 kD, which is equal to the 146 amino acid sequence carboxy terminal to the signal sequence. The recombinant *ob* protein is not modified when expressed in Pichia pastoris. Expression of mammalian genes in Pichia generally results in the formation of the correct protein structure[Cregg, 1993 #792]. These findings suggest that the *ob* protein is not glycosylated and is not post-translationally processed *in vivo*. The data do not exclude the possibility that the *ob* protein is noncovalently bound to itself or other proteins in plasma or adipose tissue. Although proteolytic cleavage of the protein has not been excluded, lower molecular weight forms of the *ob* protein were not detected by any of the antisera used, including four anti-peptide antibodies.

The *ob* protein has two cysteine residues and circulates as a monomer in human, and as a monomer and dimer in mouse. An intramolecular disulphide bond typical of secreted molecules is found when the human protein is expressed in Pichia pastoris suggesting that it is likely to be present *in vivo*. This is supported by the bioactivity of the recombinant bacterial protein, which has an intramolecular disulphide bond. The mouse *ob* protein can be found in plasma as a monomer and as a dimer. That monomer and dimer are seen when the mouse *ob* protein is expressed in yeast shows that the propensity of the mouse protein to form a dimer is a result of differences in the primary sequence relative to human. While it is clear that the monomer has bioactivity, the functional activity of the dimer is unknown.

The effect of the *ob* protein on food intake and body weight in *ob* mice is dramatic. After three weeks treatment, the *ob* mice receiving daily injections of recombinant protein had lost 40% of their weight and were consuming 40 % as much food as control animals. Moreover, the weight of the treated *ob* mice had not yet equilibrated at the time the experiment was terminated. The results of the pair feeding experiment indicate weight loss is a result of effects on both food intake and energy expenditure. Thus, a separate group of *ob* mice whose caloric intake was restricted to that of *ob* mice receiving protein lost significantly

less weight than the animals receiving protein. The reduction in food intake in ob/ob mice to a level lower than that of wild type mice, within a day of receiving the *ob* protein, indicates that they are especially sensitive to its effects. Indeed, the *ob* receptor may be upregulated in these animals. Food intake of treated *ob* mice became relatively constant after five days of treatment. If this is the result of the protein having reached steady state levels, it would suggest that the protein has a relatively long half life[Goodman, 1990 #793]. This conclusion is consistent with data from parabiosis experiments[Coleman, 1978 #42; Weigle, 1988 #349].

Effects of recombinant protein on the body weight of wild type mice were small but statistically significant during the first two weeks of the study. While the difference in weight between wild type mice receiving protein vs. PBS was sustained at later time points, the statistical significance of the data had greatly diminished after three weeks. The early weight loss could not be accounted for by a difference in food intake. Presumably, the measurement of food intake was not precise enough to detect a decrease resulting in a one gram difference in body weight during treatment. These observations differ from the results of previous experiments in which wild type rodents have been joined by parabiotic union to *db* mice, *fatty* rats, rats with hypothalamic lesions and rats rendered obese by a high calorie diet[Coleman, 1978 #42; Harris, 1987 #800; Harris, 1989 #799; Hervey, 1959 #305]. In each case, the wild type animals become anorectic and lose copious amounts of weight. As the levels of *ob* protein are increased in *db* mice and *fa* rats and the level of *ob* RNA is increased in mice with hypothalamic lesions, it is likely that wild type mice can respond to *ob* when it circulates in plasma at a sufficiently high level. The findings reported here are consistent with the possibility that the levels of the administered protein were below endogenous levels, leading to equilibration at a slightly lower body weight. Quantitation of the circulating levels of the *ob* protein in the treated mice will resolve this issue. Immunoprecipitations have suggested that the levels of circulating *ob* protein were not substantially elevated in the wild type mice receiving protein

The lesser effect of the protein on wild type mice and the absence of a response in *db* mice makes it unlikely that the treatment has nonspecific or aversive effects. All of the *db* mice lost a small amount of weight during the treatment period, whether or not they were receiving the *ob* protein. The *db* animals were markedly hyperglycemic and the weight loss is likely to be the result of diabetes and not the experimental protocol. C57BL/Ks db/db mice often develop diabetes and begin to lose small amounts of weight when of the age of

the animals used in this study[Coleman, 1978 #42]. C57Bl/6J ob/ob mice of a similar age do not develop significant hyperglycemia. These phenotypic differences are thought to be the result of genetic differences in the strains (C57Bl6J vs. C57Bl/Ks) carrying the mutations.[Coleman, 1978 #42].

The failure to detect the truncated 105 amino acid protein predicted by the cDNA sequence of the *ob* gene in C57Bl/6J ob/ob mice suggests that the mutant protein is either degraded or not translated. However, the possibility that the antisera used do not detect this truncated protein cannot be excluded. The observed ten-fold increase in the levels of the *ob* protein in *db* mice compared to wild type suggests that the *ob* protein is overproduced when there is resistance to its effects. These data correlate with studies of the *ob* mRNA (Next Example) . As mentioned, previous experiments have shown that mutations of the mouse *db* and the rat *fa* genes, which map to homologous chromosomal regions, result in overproduction of a plasma factor that suppresses body weight[Truett, 1991 #409; Coleman, 1978 #42; Hervey, 1959 #305]. In both cases, it has been suggested that the mutant animals are resistant to the effects of the *ob* protein. This possibility is confirmed by the observation that the *ob* protein has no effect on body weight or food intake when administered to *db* mice.

Obesity in humans could be associated with increased levels of the *ob* protein in plasma in individuals who are relatively unresponsive to the hormone. On the other hand, reduced expression of *ob* could also lead to obesity in which case "normal" (i.e; inappropriately low) levels of the protein might be found. Thus, the levels of *ob* protein in human plasma could be a marker for different forms of obesity. In a small group of lean subjects with BMI <25, low nanogram levels of circulating *ob* protein are detectable by ELISA. Significantly, variable concentrations were noted suggesting that the level of expression and/or sensitivity to the protein may play a role in determining body weight.

The site of action of the *ob* protein is unknown. The protein affects both food intake and energy expenditure, a finding consistent with clinical studies indicating that alterations of both systems act to regulate body weight[Leibel, 1995 #795; Keesey, 1984 #796]. The hypothalamus is likely to be downstream of *ob* in the pathway that controls body weight, although direct effects on a variety of organs are possible

References for this example

1. Kennedy, G.C. The role of depot fat in the hypothalamic control of food intake in the rat. Proc. Roy. Soc. (London) (B)., 1953. 140: 578-592.

2. Leibelt, R.A., S. Ichinoe, and N. Nicholson. Regulatory influences of adipose tissue on food intake and body weight. Ann. N.Y. Acad. Sci., 1965. 131: 559-582.

3. Zhang, Y., P. Proenca, M. Maffei, M. Barone, L. Leopold, and J.M. Friedman. Positional cloning of the mouse *obese* gene and its human homologue. Nature., 1994. 372: 425-432.

4. Maffei, M., H. Fei, G.W. Lee, C. Dani, P. Leroy, Y. Zhang, R. Proenca, R. Negrel, G. Ailhaud, and J.M. Friedman. Increased expression in adipocytes of *ob* RNA in mice with lesions of the hypothalamus and with mutations at the *db* locus. (Next Example)

5. Bahary, N., R.L. Leibel, L. Joseph, and J.M. Friedman. Molecular mapping of the mouse db mutation. Proc. Nat. Acad. Sci. USA., 1990. 87: 8642-8646.

6. Truett, G.E., N. Bahary, J.M. Friedman, and R.L. Leibel. The rat obesity gene fatty *(fa)* maps to chromosome 5: Evidence for homology with the mouse gene diabetes *(db)*. Proc. Natl. Acad. Sci. USA., 1991. 88: 7806-7809.

7. Abrams, J.S., M.-G. Roncarolo, H. Yssel, U. Andersson, G.J. Gleich, and J.E. Silver. Strategies of anti-cytokine monoclonal antibody development: Immunoassay of IL-10 and IL-5 in clinical samples. Immunological Reviews., 1992. 5-24.

8. Ellman, G.L. Arch. Biochem. Biophy., 1959. 82: 70-77.

9. Coleman, D.L. Obese and Diabetes: two mutant genes causing diabetes-obesity syndromes in mice. Diabetologia., 1978. 14: 141-148.

10. Cregg, J.M., T.S. Vedvick, and W.C. Raschke. Recent advances in the expression of foreign genes in *Pichia pastoris*.. Bio/Technology., 1993. 11: 905-914.

11. Goodman, L.S., , in *The Pharmacological Basis of Therapeutics.*, A. Gilman, 1990, Pergamon Press: New York. p. 19-45.

12. Weigle, D.S. Contribution of decreased body mass to diminished termic/effect of exercise in reduced-obese men. Int. J. Obesity., 1988. 12: 567-578.

13. Harris, R.B.S., E. Hervey, G.R. Hervey, and G. Tobin. Body composition of lean and obese zucker rats in parabiosis. Int. J. Obes., 1987. 11: 275-283.

14. Harris, R.B.S. and R.J. Martin, Physiological and metabolic changes in parabiotic partners of obese rats., in *Hormones, Thermogenesis and Obesity,* H. Lardy and F. Straatman, 1989, Elsevier Science Publishing Co.: New York.

15. Hervey, G.R. The effects of lesions in the hypothalamus in parabiotic rats. J. Physiol., 1959. 145: 336-352.

16. Leibel, R.L., M. Rosenbaum, and J. Hirsch. Changes in energy expenditure resulting from altered body weight. The New England Journal of Medicine., 1995. 332: 621-628.

17. Keesey, R.C. and S.W. Corbett, Metabolic defense of the body weight set-point, in *Association for Research in Nervous and Mental Disease,* A.J. Stunkard and E. Stellar, 1984, Raven Press: New York. p. 87-96.

EXAMPLE: Increased Expression in Adipocytes of Mice of *ob* mRNA Due to Lesions of Hypothalamus and with Mutations at the *db* Locus

(Various references are cited by author, year, and # in this Example, which citations correlate with the list of references found at the end of this Example.)

The gene product of the recently cloned mouse obese gene (*ob*) plays an important role in regulating the adipose tissue mass. *ob* RNA is expressed specifically by mouse adipocytes in vivo in each of several different fat cell depots including brown fat. It is also expressed in cultured 3T3-442A preadipocyte cells that have been induced to differentiate. Mice with lesions of the hypothalamus, as well as mice mutant at the *db* locus, express a twenty-fold higher level of *ob* RNA in adipose tissue. These data suggest that both the *db* gene and the hypothalamus are downstream of the *ob* gene in the pathway that regulates the adipose tissue mass and are consistent with previous experiments suggesting that the *db* locus encodes the *ob* receptor. In the db/db and lesioned mice, quantitative differences in the level of expression of *ob* RNA correlated with the lipid content of adipocytes. The molecules that regulate the level of expression of the *ob* gene in adipocytes are likely to play an important role in determining body weight as are the molecules that mediate the effects of *ob* at its site of action.

MATERIALS AND METHODS

In Situ Hybridization

White fat tissues from identical abdominal regions of wild type (wt) and *db* mice were processed simultaneously according to the modified ethod described by Richardson et al[6]. Briefly, tissues were fixed in Bouin's solution for 2 hours at 4°C. They were then dehydrated by serial treatment of increasing

concentrations of ethanol from 10% to 100%, each for 5 min. at 4°C. Further incubation of tissues with xylene (1h) and paraffin (2h) were performed at 65°C. Embedded wt and db/db fat tissues were sectioned and mounted on to the samme conditions later. Sections were baked at 65°C for 1h and treated with xylene and serial dilutions of ethanol from 100% to 50%, each for 3 min. at room temperature. Antisense RNA probe of ob gene was synthesized by in vitro transcription of linearized ob gene coding sequence upstream of a Sp6 RNA polymerase promoter. In situ hybridization was carried out exactly according to Schaeren-Wiemers and Gerfin-Moser[7].

RNA Preparation and Cell Culture

Total RNA and Northern blots were prepared as described[3].

Stromal vascular cells and adipocytes were prepared according to Rodbell and RNA from both fractions was prepared according to Dani et al[8, 9]. After sub-cloning, 3T3-F442 cells were grown in Dulbecco's modified Eagle medium containing 10% foetal bovine serum (defined as standard medium) [10]. At confluence, cells were treated in standard medium supplemented with 2 nM triiodothyronine (T3) and 17 nM insulin. Twelve days later, RNA was prepared as above.

Gold ThioGlucose Treatment

Two month old female CBA/J mice were treated with a single intraperitoneal injection of aurothioglucose (Sigma A0632) at a dose of .2 mg/g in normal saline. Control animals were injected with normal saline. Mice were weighed one month after the treatment. Adipose tissue RNA was isolated from those treated animals whose weight had increased more that twenty grams post GTG treatment.

RESULTS

The *ob* gene was recently found to be expressed in adipose tissue[3]. As adipose tissue is composed of many cell types including adipocytes, preadipocytes, fibroblasts and vascular cells, in situ hybridization was performed to sections of epididymal fat pads from normal animals with sense and antisense *ob* riboprobes[6, 11]. When using the antisense probe, positive signals were detectable in all of the adipocytes in the section (Fig. 29-labeled Wt). Signals were not noted when the antisense probe was hybridized to sections of brain (data not shown). Hybridization of the antisense probe to sections of adipose tissue from C57Bl/Ks db/db mice was greatly increased, confirming the adipocyte specific expression of *ob* RNA and demonstrating a large increase in the level of *ob* RNA per adipocyte in these animals (Fig. 29-labeled db/db). Mice mutant at the *db* locus are massively obese as part of a syndrome that is phenotypically identical to that seen in C57Bl/6J ob/ob mice[12].

ob RNA was not synthesized by adipose tissue stromal cells separated from adipocytes. As expected, cells in the adipocyte fraction expressed *ob* RNA using Northrn blots (Fig. 30). The same result was obtained using RT-PCR (data not shown). These data support the conclusion that only adipocytes express the *ob* gene. Data from cultured adipocytes confirm this conclusion. In these studies, 3T3-F442A cells were cultured using conditions that lead to lipid accumilation, as part of a cellular program leading to differentication into adipocytes. *ob* RNA was not expressed in exponentially growing cells as well as in confluent 3T3-F442A preadipocyte cells which express early markers while differentiation of these cells into adipocytes led to the expression of defectable levels of ob RNA (Fig. 30)[13]. The level of ob RNA is extremely

sensitive to the culture conditions as no message was observed in late post-confluent cells not exposed to insulin.

Hybridization studies showed that ob RNA is expressed *in vivo* in several different fat depots including the epididymal, parametrial, abdominal, perirenal, and inguinal fat pads (Fig. 3IA). The precise level of expression in each of the depots was somewhat variable, with inguinal and parametrial fat expressing lower levels of *ob* RNA. *ob* RNA is also expressed in brown adipose tissue although the level of expression is ~ 50 fold lower in brown fat relative to the other adipose tissue depots. These quantitative differences correlated loosely with previously reported differences in cell size among the different fat cell depots[14]. The amount of *ob* RNA in brown fat is unaffected by cold exposure (Fig. 3IB). In this experiment, the level of uncoupling protein RNA (UCP) increased in brown fat after cold exposure while the level of *ob* RNA did not change[15]. In aggregate, these data confirm that all adipocytes are capable of producing *ob* RNA and demonstrate a variable level of expression in different fat depots. These data support the possibility that the level of the encoded protein correlates with the total adipose tissue mass.

We next measured the levels of *ob* RNA in db/db mice and mice with lesions of the hypothalamus. Lesions of the ventromedial hypothalamus (VMH) result in obesity as part of a syndrome resembling that seen in ob/ob and db/db mice[16]. Parabiosis experiments suggest such lesions result in over expression of a blood borne factor that suppresses food intake and body weight[17]. Similar results are noted when mice mutant at the *db* locus are parabiosed to normal mice, suggesting the *ob* receptor may be encoded by the *db* locus[18]. Thus, obesity resulting from VMH lesions and the *db* mutation

may be the result of resistance to the effects of the *ob* protein. If so, a secondary increase in the levels of *ob* RNA in adipose tissue would be predicted.

Hypothalamic lesions were induced in female CBA mice using the chemical Gold ThioGlucose (GTG)[19]. This treatment results in specific hypothalamic lesions, principally in the ventromedial hypothalamus (VMH), with the subsequent development of obesity within several weeks (manuscript in preparation). In our experience, a single intraperitoneal injection of GTG of .2 mg/gm body weight results in the development of obesity within four weeks. One month old female CBA/J mice (20-25 grams) were treated with GTG and the subsequent weight gain of treated and control animals is shown. (Table 2) Adipose tissue RNA was prepared from db/db mice and from those GTG treated animals that gained >20 gm. Northern blots showed a twenty-fold increase in the level of *ob* RNA in two month old db/db and GTG treated mice compared to normal animals (Fig. 3).

TABLE 2. Weight Gain in Gold ThioGlucose Treated Mice

	control (n = 41)	GTG (n = 93)
<10 g	41, (100%)	4, (4%)
10 g - 20 g	0, (0%)	15, (16%)
>20 g	0, (0%)	74, (80%)

Two month old female CBA/J mice were treated with goldthioglucose (GTG). Goldthioglucose (Sigma A0632) was administered intraperitonealy in normal saline solution at a dosage of 2.0 mg/g. Body weight of control and injected animals was recorded before and one month after the injection. Animals were housed five to a cage and were fed *ad libitum*. The amount of weight gained one month postinjection is shown in the table. Animals with a body weight gain greater that 20 g one month after injection were selected for further study.

The gene product of the mouse *ob* gene circulates in mouse and human plasma where it may act to regulate the adipose tissue mass (manuscript in preparation). Further studies on the regulation of expression and mechanism of action of *ob* will have important implications for our understanding of the physiologic pathway that regulates body weight.

In this report we show that the *ob* gene product is expressed exclusively by adipocytes in all adipose tissue depots. This result is consistent with the possibility that the protein product of the *ob* gene correlates with the bodies lipid stores. Moreover *ob* RNA is upregulated twenty fold in *db* mice and mice with hypothalamic lesions. In these animals, the actual increase in the level of *ob* RNA per cell is likely to be even higher than twenty fold since the adipocyte cell size is increased ~five fold in these animals (see Fig. 29) [14]. These data position the *db* gene and the hypothalamus downstream of *ob* in the pathway that controls body weight and is consistent with the hypothesis that the *ob* receptor is encoded at the *db* locus[18]. The molecular cloning of the *ob* receptor and/or the *db* gene will resolve this issue. The increase in the level of *ob* RNA in db/db and GTG treated mice also suggests a non cell-autonomous function of the ob gene product in fat cells[4, 5]. Thus, if the encoded protein acted directly on fat cells to inhibit growth or differentiation, the overexpression of the wild type *ob* gene in GTG treated mice would result in a lean phenotype.

The most parsimonious explanation of these data is that the *ob* protein functions as an endocrine signaling molecule that is secreted by adipocytes and acts, directly or indirectly, on the hypothalamus. Direct effects on the hypothalamus would require that mechanisms exist to allow passage of the *ob* gene product across the blood brain barrier. Mechanisms involving the circumventricular organ and/or specific transporters could permit brain access of a molecule the size of that encoded by the *ob* gene[20-22]. However, this

hypothesis must be considered with caution until the means by which the protein might cross the blood brain barrier have been identified. Moreover, possible effects on other target organs will need to be evaluated.

The fat cell signal(s) that are responsible for the quantitative variation in the expression level of the *ob* gene is not yet known but correlates with differences in adipocyte cell size. Adipocytes from db/db mice are five times as large as those from normal mice, with a cell size of ~1.0 µg lipid/cell[14]. Prior evidence has indicated that fat cell lipid content and/or size is an important parameter in determining body weight[23, 24]. It could be that each fat cell expresses a low level of *ob* RNA that further increases in proportion to the cell size. It is also possible that cell size is not the sensed parameter and merely correlates with the intracellular signal that increases the expression of the *ob* gene in adipocytes from db/db and VMH lesioned mice. In any case, the components of the signal transduction pathway regulating the synthesis of *ob* RNA are likely to be important in determining body weight. Genetic and environmental influences that reduce the level of expression of *ob* would act to increase body weight as would influences that decreased sensitivity to the encoded protein. The specific molecules that regulate the level of expression levels of the *ob* gene are as yet unknown, and await a determination of the level(s) of gene control that leads to quantitative variation in the level of *ob* RNA and an examination of the regulatory elements of the *ob* gene. The identification of the molecules that regulate the expression of the *ob* gene in adipocytes and those that mediate the effects of the encoded protein at its site(s) of action will greatly enhance our understanding of the physiologic mechanisms that regulate body weight..

1. Kennedy, G.C. The role of depot fat in the hypothalamic control of food intake in the rat. Proc. Roy. Soc. (London) (B), 1953. 140: 578-592.

2. Leibelt, R.A., S. Ichinoe, and N. Nicholson. Regulatory influences of adipose tissue on food intake and body weight. Ann. N.Y. Acad. Sci, 1965. 131: 559-582.

3. Zhang, Y., P. Proenca, M. Maffei, M. Barone, L. Leopold, and J.M. Friedman. Positional cloning of the mouse *obese* gene and its human homologue. Nature, 1994. 372: 425-432.

4. Ashwell, M., C.J. Meade, P. Medawar, and C. Sowter. Adipose tissue: contributions of nature and nurture to the obesity of an obese mutant mouse (ob/ob). Proc. R. Soc. Lond., 1977. 195: 343-353.

5. Ashwell, M. and C.J. Meade. Obesity: Do fat cells from genetically obese mice (C57BL/6J ob/ob) have an innate capacity for increased fat storage? Diabetologia, 1978. 15: 465-470.

6. Richardson, R.L., J.T. Wright, J.-W. Kim, and G.J. Hausman. Expression of transforming growth factor-β (TGF-β1) and insulin-like growth factor II (IGF-II) messenger RNA in the developing subcutaneous tissue (SQ) of the fetal pig. Growth, Development & Aging, 1992. 56: 149-157.

7. Schaeren, N. and A. Gerfin-Moser. A single protocol to detect transcripts of various types and expression levels in neural tissue and cultural cells: *in situ* hybridization using digoxigenin-labeled cRNA probes. Histochemistry, 1993. 100: 431-440.

8. Rodbell, M. Metabolism of isolated fat cells. J. Biological Chemistry, 1964. 239: 375-380.

9. Dani, C., A. Doglio, A. Pradines-Figueres, and P. Grimaldi, Molecular biology techniques in the study of adipocyte differentiation., in *Obesity in Europe: 88,* P. Bjorntorp and R. Rossner, Editor^Editors. 1989, John Libbey Company Ltd.: London, England. p. 371-376.

10. Dani, C., et al. Regulation of gene expression by insulin in adipose cells: opposite effects on adipsin and glycerophosphate dehydrogenase genes. Mol. Cell. Endocrinol., 1989. 63: 199-208.

11. Wasserman, M., The concept of the fat organ: in Rodahl, Issekutz, fat as a tissue, in *Rodahl, Issekutz, Fat as a tissue.* 1964, McGraw Hill: New York. p. 22-92.

12. Bahary, N., R.L. Leibel, L. Joseph, and J.M. Friedman. Molecular mapping of the mouse db mutation. Proc. Nat. Acad. Sci. USA, 1990. 87: 8642-8646.

13. Dani, C.A., A. Doglio, E.Z. Amri, S. Bardon, P. Fort, B. Bertrand, P. Grimaldi, and G. Ailhaud. Cloning and regulation of a mRNA specifically expressed in the preadipse state. J. Biol. Chem., 1989. 264: 10119-10125.

14. Johnson, P.R. and J. Hirsch. Cellularity of adipose depots in six strains of genetically obese mice. Journal of Lipid Research, 1972. 13: 2-11.

15. Jacobsson, A., U. Stadker, M.A. Glotzer, and L.P. Kozak. Mitochondrial uncoupling protein from mouse brown fat. J. Biol. Chem., 1985. 260: 16250-16254.

16. Bray, G.A. and L.A. Campfield. Metabolic factors in the control of energy stores. Metabolism, 1975. 24: 99-117.

17. Hervey, G.R. The effects of lesions in the hypothalamus in parabiotic rats. J. Physiol., 1959. 145: 336-352.

18. Coleman, D.L. Obese and Diabetes: two mutant genes causing diabetes-obesity syndromes in mice. Diabetologia, 1978. 14: 141-148.

19. Debons, A.F., I. Krimsky, M.L. Maayan, K. Fani, and F.A. Jimenez. Gold thioglucose obesity syndrome. Fed. Proc., 1977. 36: 143-147.

20. Johnson, A.K. and P.M. Gross. Sensory circumventricular organs and brain homeostatic pathways. FASEB Journal, 1983. 7: 678-686.

21. Baura, G.D., D.M. Foster, D. Porte Jr., S.E. Kahn, R.N. Bergman, C. Cobelli, and M.W. Schwartz. Saturable transport of insulin from plasma into the central nervous system of dogs in vivo. Jr. of Clinical Investigation, Inc., 1993. 92: 1824-1830.

22. Pardridge, W.M. Receptor-mediated peptide transport through the blood-brain barrier. Endocrine Reviews, 1986. 7: 314-330.

23. Faust, I.M., P.R. Johnson, J.S. Stern, and J. Hirsch. Diet-induced adipocyte number increase in adult rats: a new model of obesity. Am. Journal of Physiol., 1978. 235: E279-86.

24. Faust, I.M., P.R. Johnson, and J. Hirsch. Surgical removal of adipose tissue alters feeding behavior and the development of obesity in rats. Science, 1977. 197: 393-396.

THE REMAINDER OF THIS PAGE IS INTENTIONALLY LEFT BLANK

EXAMPLE: RNA Expression Pattern and Mapping of the Human *ob* Gene on the Physical, Cytogenetic, and Genetic Maps of Chromosome 7

(Various references are cited by author, year, and # in this Example, which citations correlate with the list of references found at the end of this Example.)

OB RNA is expressed at high levels in human adipose tissue and at substantially lower levels in placenta and heart. The human OB gene maps to a large YAC contig derived from chromosome 7q31.3. In addition to confirming the relative location of the gene based on mouse-human comparative mapping, this study has identified 8 established microsatellite markers in close physical proximity to the human OB gene. Since mutations in mouse ob can result in a syndrome that closely resembles morbid obesity in humans, these genetic markers represent important tools for studying the possible role of the OB gene in inherited forms of human obesity.

ABBREVIATIONS

FISH: fluorescence in situ hybridization; GDB: Genome Data Base; kb: kilobase pairs; Mb: megabase pairs; PCR: polymerase chain reaction; STS: sequence-tagged site; YAC: yeast artificial chromosome.

MATERIAL AND METHODS

Northern blot analysis

Total RNA was prepared from adipose tissue using the method of Chirgwin et al. (1979). Northern blots, radiolabelling, and hybridizations were performed as described (Zhang et al. 1994). Northern blots of polyA+ RNA [human MTN, human MTN II, and human fetal MTN II] were purchased from CLONTECH (Palo Alto, CA), as were PCR primers used to generate the radiolabelled human actin probe.

STS development

STS-specific PCR assays were developed and optimized essentially as described (Green and Green, 1991a; Green et al. 1991b; Green, 1993; Green et al. 1994). Each STS is named using the prefix sWSS followed by a unique number. Details about the 19 STSs reported here are provided in Table 1, with additional information (e.g., PCR reaction conditions, complete DNA sequence) available in GenBank and/or the Genome Data Base (GDB). For the microsatellite-specific STSs, the oligonucleotide primers used in the PCR assays (Table 3) corresponded either to those employed for genotype analysis (Table 4) or those designed [most often with the computer program OSP (Hillier and Green, 1991)] using the DNA sequence available in GenBank.

Table 3. STSs in the YAC contig containing the human OB gene

The 19 chromosome 7-specific STSs mapped to the YAC contig containing the human OB gene (Figure 34) are listed. In each case, the designated sWSS name, relevant alias, GDB-assigned locus name, STS source, PCR primer sequences, STS size, and GDB identification number are indicated. The sources of STSs are as follows: YAC End [isolated insert end of a YAC (Green, 1993)], Lambda Clone [random chromosome 7-specific lambda clone (Green et al. 1991b; Green, 1993)], Genetic Marker [microsatellite marker (Green et al. 1994), see Table 2], YAC Insert [random segment from YAC insert], and Gene [gene-specific STS]. Note that for some genetic marker-specific STSs, the PCR primers used for identifying YACs (listed in this table) are different from those used for performing genotype analysis (Table 4), since the detection of YACs containing a genetic marker does not require amplification of the polymorphic tract itself. All of the indicated PCR assays utilized an annealing temperature of 55oC, except for sWSS494, sWSS883, sWSS1529, and sWSS2619 (which used 50oC), sWSS999 and sWSS1174 (which used 60oC), and sWSS808 (which used 65oC). Additional details regarding the STS-specific PCR assays are available in GDB.

STS Name	Alias	Locus	Source	PCR Primers	Size (bp)	GDB ID N
sWSS1734		D7S2185	YAC End	CAAGACAAATGAGATAAGG AGAGTTACAGCTTTACAG	72	G00-455-235
sWSS494		D7S2016	Lambda Clone	CTAAACACCTTTCCATTCC TTATATTCACTTTTCCCCTCTC	112	G00-334-404
sWSS883	UT528	D7S1498	Genetic Marker	TGCAGTAAGCTGTGATTGAG GTGCAGCTTTAATTGTGAGC	490	G00-455-262
sWSS2359	AFMa065zg9	D7S1873	Genetic Marker	AGTGTTGTGTTTCTCCTG AAAGGGGATGTGATAAGTG	142	G00-455-247
sWSS2336	AFMa125wh1	D7S1874	Genetic Marker	GGTGTTACGTTTAGTTAC GGAATAATGAGAGAAGATTG	112	G00-455-244
sWSS1218	AFM309yf1	D7S680	Genetic Marker	GCTCAACTGACAGAAAAC GACTATGTAAAAGAAATGCC	154	G00-307-733
sWSS1402		D7S1916	YAC End	AAAGGGCTTCTAATCTAC CCTTCCAACTTCTTTGAC	137	G00-344-044
sWSS999		D7S1674	YAC Insert	TAAACCCCCTTTCTGTTC TTGCATAATAGTCACACCC	105	G00-334-839
sWSS1751		D7S2186	YAC End	CCAAAATCAGAATTGTCAGAAG AAACCGAAGTTCAGATACAG	186	G00-455-238
sWSS1174	AFM218xf10	D7S514	Genetic Marker	AATATCTGACATTGGCAC TTAGACCTGAGAAAAGAG	144	G00-307-700
sWSS2061		D7S2184	YAC End	GTTGCACAATACAAAATCC CTTCCATTAGTGTCTTATAG	200	G00-455-241

sWSS2588		D7S2187	YAC End	ATCACTACACACCTAATC CCATTCTACATTTCCACC	117	G00-455-253
sWSS808	PAX4	PAX4	Gene	GGCTGTGTGAGCAAGATCCTAGGA TTGCCAGGCAAAGAGGGCTGGAC	153	G00-455-2
sWSS1392	AFM206xc1	D7S635	Genetic Marker	CTCAGGTATGTCTTTATC TGTCTCTGCATTCTTTTC	75	G00-307-815
sWSS1148	AFM199xh12	D7S504	Genetic Marker	GACACATACAAACACAAG ATTGAGTTGAGTGTAGTAG	60	G00-307-652
sWSS1529		D7S1943	YAC End	CAGGGATTTCTAATTGTC AAAAGATGGAGGCTTTTG	116	G00-334-119
sWSS2619	OB G00-455-256	OB	Gene	CGTTAAGGGAAGGAACTCTGG TGGCTTAGAGGAGTCAGGGA	106	
sWSS404		D7S1956	Lambda Clone	ACCAGGGTCAATACAAAG TAATGTGTCCTTCTTGCC	122	G00-334-251
sWSS2367	AFMa345wc9	D7S1875	Genetic Marker	CAATCCTGGCTTCATTTG AAGGTGGGTAGGATGCTA	81	G00-455-250

Table 4. Microsatellite markers in the YAC contig containing the human OB gene

The 8 microsatellite markers mapped to the YAC contig containing the human OB gene (Figure 3) are listed. In each case, the marker name (indicated as the alias in Table 3), type of microsatellite motif [tetranucleotide (Tetra) repeat or (CA)n repeat], GDB-assigned locus name, primer sequences utilized for PCR-based genotype analysis, and GDB identification number are indicated. Additional details regarding the PCR assays and the polymorphisms are available in GDB.

Marker Name	Type	Locus	Primers	GDB ID No.
UT528	Tetra.	D7S1498	TGCAGTAAGCTGTGATTGAG GTGCAGCTTTAATTGTGAGC	G00-312-446
AFMa065zg9	(CA)n	D7S1873	AGCTTCAAGACTTTNAGCCT GGTCAGCAGCACTGTGATT	G00-437-253
AFMa125wh1	(CA)n	D7S1874	TCACCTTGAGATTCCATCC AACACCGTGGTCTTATCAAA	G00-437-263
AFM309yf10	(CA)n	D7S680	CATCCAAGTTGGCAGTTTTT AGATGCTGAATTCCCAGACA	G00-200-283
AFM218xf10	(CA)n	D7S514	TGGGCAACACAGCAAA TGCAGTTAGTGCCAATGTCA	G00-188-404
AFM206xc1	(CA)n	D7S635	CCAGGCCATGTGGAAC AGTTCTTGGCTTGCGTCAGT	G00-199-240
AFM199xh12	(CA)n	D7S504	TCTGATTGCTGGCTGC GCGCGTGTGTATGTGAG	G00-188-280
AFMa345wc9	(CA)n	D7S1875	AGCTCTTGGCAAACTCACAT GCCTAAGGGAATGAGACACA	G00-437-259

The human OB-specific STS (sWSS2619) was designed using DNA sequence obtained from the 3'untranslated region of the cDNA. The human PAX4-specific STS (sWSS808) was developed using the following strategy. Oligonucleotide primers specific for the mouse Pax4 gene [GGCTGTGTGAGCAAGATCCTAGGA and GGGAGCCTTGTCCTGGGTACAAAG (Walther et al. 1991)] were used to amplify a 204-bp fragment from human genomic DNA (which was the same size product as that generated from mouse genomic DNA). This PCR assay was not suitable for identifying corresponding YACs, since a similarly-sized (200-bp) product was also amplified from yeast DNA. However, DNA sequence analysis of the PCR product generated from human DNA revealed substitutions at 20 positions among the 156 bases analyzed (data not shown). Using this human-specific sequence, a new primer (TTGCCAGGCAAAGAGGGCTGGAC) was designed and used with the first of the above mouse Pax4-specific primers (see Table 3). The resulting human PAX4-specific PCR assay did not amplify a significant product from yeast DNA and was thus used for identifying corresponding YACs.

Identification of YACs by PCR-based screening

Most of the YACs depicted in Figure 4 were derived from a collection of clones highly enriched for human chromosome 7 DNA [the chromosome 7 YAC resource (Green et al. 1995a)] using a PCR-based screening strategy (Green et al. 1995a; Green and Olson, 1990). In a few cases, clones were isolated by PCR-based screening (Green and Olson, 1990) of available total human genomic YAC libraries constructed at CEPH (Dausset et al. 1992; Albertsen et al. 1990) or ICI (Anand et al. 1990; Anand et al. 1989). Each YAC is named using the prefix yWSS followed by a unique number.

RESULTS AND DISCUSSION

Examination of the tissue expression of the human OB gene by northern blot analysis revealed that OB RNA is expressed at a high level in human adipose tissue and much lower levels in placenta and heart (Figure 3). The size of the RNA (~4.5 kb) was equivalent in human and mouse as well as in each of the expressing tissues. In these studies, five-fold higher signals were seen in 10 μg of total adipose tissue RNA as in 2 μg of polyA+ placental RNA. A five-fold lower signal was seen in polyA+ RNA from heart compared to placenta. It is estimated that the level of OB RNA is ~250-fold lower in placenta than in adipose tissue. In this experiment, OB RNA was not detected in any of the other tissues analyzed, including brain, lung, liver, skeletal muscle, kidney,

and pancreas. Additional experiments did not reveal OB RNA in spleen, thymus, prostate, testis, ovary, small intestine, colon, peripheral blood leukocytes, or in fetal brain, liver, or kidneys (data not shown). It is possible that OB is expressed at an undetectable level (by northern blot analysis) in these latter tissues or in other tissues that were not studied. The observed pattern of expression in human differs somewhat from mouse, in which ob RNA is detected almost exclusively in adipose tissue.

Comparative mapping of the ob gene region in the mouse and human genomes

The mouse ob gene is located on proximal chromosome 6 in a region homologous with a portion of human chromosome 7q. Genes within this segment include (from proximal to distal): Met protooncogene, the cystic fibrosis transmembrane conductance regulator (Cftr), paired box-containing gene 4 (Pax4), ob, and carboxypeptidase A (Cpa) (Zhang et al. 1994; Friedman et al. 1991). In mouse, genetic mapping was used to demonstrate that Pax4 is tightly linked to ob (Walther et al. 1991; Zhang et al. 1994). The physical distance between ob and Pax4 was found to be ~1 megabase pairs (Mb) (Zhang et al. 1994). Based on these comparative mapping studies, it was expected that the human OB gene would reside between PAX4 and CPA on chromosome 7q. Furthermore, since human CFTR (Heng et al. 1993) and PAX4 (Tamura et al. 1994) were mapped by fluorescence in situ hybridization (FISH) to 7q31.3 and 7q32, respectively, the most likely cytogenetic position of the human OB gene would be in the vicinity of the 7q31.3-q32 boundary.

Mapping the OB gene on human chromosome 7

An STS (sWSS2619) amplifying a small segment of the 3 untranslated region of the human OB gene was used to screen a collection of YAC clones that is highly enriched for human chromosome 7 DNA (Green et al. 1995a), and 9 YACs were identified (yWSS691, yWSS1332, yWSS1998, yWSS2087, yWSS3319, yWSS3512, yWSS4875, yWSS4970, and yWSS5004). To verify that these YACs contain the authentic human OB gene, 2 additional experiments were performed. First, each of the YACs was tested with a second human OB-specific PCR assay, and all were found to be positive (data not shown). Second, yeast DNA from each clone was digested with EcoRI and analyzed by gel-transfer hybridization using a human OB cDNA-derived probe. In all instances, a single hybridizing band was seen, and this band was the same size in the YACs and a P1 clone known to contain the human OB gene (data not shown).

Using the computer program SEGMAP (Green and Green, 1991a; C.L. Magness and P. Green, unpublished data) and other YAC-based STS-content data that we have generated for chromosome 7 (Green et al. 1991b; Green et al. 1994; Green et al. 1995a), the human OB gene

was found to reside within the YAC contig depicted in Figure 2. Specifically, this contig consists of 43 overlapping YACs and 19 uniquely-ordered STSs. Details about each of the 19 STSs are provided in Table 3. In addition to the OB-specific STS, the contig also contains an STS (sWSS808) specific for the human PAX4 gene (Tamura et al. 1994; Stapleton et al. 1993), 7 STSs derived from chromosome 7-specific YACs, 2 STSs derived from chromosome 7-specific lambda clones, and, importantly, 8 microsatellite-specific STSs. Additional details about these 8 genetic markers, including sequences of the primers used for genotype analysis, are provided in Table 4. Of note, there is redundant YAC-based connectivity throughout the contig (i.e., there are 2 or more YACs connecting each adjacent pair of STSs), lending strong support for the relative order of STSs shown in Figure 3.

As depicted in Figure 3, the predicted orientation of the human OB-containing YAC contig is such that sWSS1734 is the centromeric-most STS (i.e., closest to CFTR) while sWSS2367 is the telomeric-most STS (i.e., closest to CPA). This orientation is predominantly based on comparative mapping data, which places Pax4 proximal and ob distal within the syntenic block present in mouse and human DNA (Zhang et al. 1994). The OB gene maps near the telomeric end of the contig, based on the placement of the OB-specific STS (sWSS2619).

While the contig shown in Figure 3 was deduced by SEGMAP without consideration of YAC sizes (thereby displaying STSs equidistant from one another), a similar analysis of the data by SEGMAP that accounted for YAC sizes indicated that the total size of the region covered by the contig is just over 2 Mb (data not shown). Thus, while all 8 of the microsatellite-specific STSs (Table 4) are contained within a genomic interval spanning roughly 2 Mb, the 3 closest to the telomeric end of the contig (sWSS1392, sWSS1148, and sWSS2367) are particularly close to the OB gene itself (perhaps within an interval as small as ~500 kb). In fact, all 3 of the latter STSs are present in at least 1 of the human OB-containing YACs. Of note, the interval between human PAX4 (sWSS808) and OB (sWSS2619) is estimated to be ~400 kb, whereas this region was predicted to span ~1 Mb in mouse (Zhang et al. 1994). Finally, 3 of the YACs within the contig (yWSS691, yWSS999, and yWSS2935) have also been analyzed by FISH, and each was found to hybridize exclusively to 7q31.3 (T. Featherstone and E.D. Green, unpublished data). One of these YACs (yWSS691) contains the OB-specific STS, while the other 2 clones contain the PAX4-specific STS. The latter results are generally consistent with the previous cytogenetic assignment of human PAX4 to 7q32 (Tamura et al. 1994). Based on these data, the human OB gene can be assigned to cytogenetic band 7q31.3.

References for this Example

Albertsen, H.M., H. Abderrahim, H.M. Cann, J. Dausset, D. Le Paslier, and D. Cohen. 1990. Construction and characterization of a yeast artificial chromosome library containing seven haploid genome equivalents. Proc. Natl. Acad. Sci. USA **87:** 4256-4260.

Anand, R., A. Villasante, and C. Tyler-Smith. 1989. Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. Nucl. Acids Res. **17:** 3425-3433.

Anand, R., J.H. Riley, R. Butler, J.C. Smith, and A.F. Markham. 1990. A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. Nucl. Acids Res. **18:** 1951-1956.

Ballabio, A. 1993. The rise and fall of positional cloning? Nature Genet. **3:** 277-279.

Chirgwin, J.J., A.E. Przbyla, R.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochem. **18:** 5294-5299.

Coleman, D.L. 1978. Obese and diabetes: two mutant genes causing diabetes-obesity syndromes in mice. Diabetologia **14:** 141-148.

Collins, F.S. 1992. Positional cloning: let's not call it reverse anymore. Nature Genet. **1:** 3-6.

Collins, F. and D. Galas. 1993. A new five-year plan for the U.S. human genome project. Science **262:** 43-46.

Collins, F.S. 1995. Positional cloning moves from perditional to traditional. Nature Genetics **9:** 347-350.

Dausset, J., P. Ougen, H. Abderrahim, A. Billault, J.-L. Sambucy, D. Cohen, and D. Le Paslier. 1992. The CEPH YAC library. Behring Inst. Mitt. **91:** 13-20.

Friedman, J.M., R.L. Leibel, D.S. Siegel, J. Walsh, and N. Bahary. 1991. Molecular mapping of the mouse ob mutation. Genomics **11:** 1054-1062.

Green, E.D. and M.V. Olson. 1990. Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **87:** 1213-1217.

Green, E.D. and P. Green. 1991a. Sequence-tagged site (STS) content mapping of human chromosomes: theoretical considerations and early experiences. PCR Methods Applic. **1:** 77-90.

Green, E.D., R.M. Mohr, J.R. Idol, M. Jones, J.M. Buckingham, L.L. Deaven, R.K. Moyzis, and M.V. Olson. 1991b. Systematic generation of sequence-tagged sites for physical mapping of human chromosomes: application to the mapping of human chromosome 7 using yeast artificial chromosomes. Genomics **11:** 548-564.

Green, E.D. 1993. Physical mapping of human chromosomes: generation of chromosome-specific sequence-tagged sites. In Methods in Molecular Genetics (Vol. 1): Gene and Chromosome Analysis (Part A). (ed. K.W. Adolph), pp. 192-210. Academic Press, Inc. San Diego.

Green, E.D., J.R. Idol, R.M. Mohr-Tidwell, V.V. Braden, D.C. Peluso, R.S. Fulton, H.F. Massa, C.L. Magness, A.M. Wilson, J. Kimura, J. Weissenbach, and B.J. Trask. 1994. Integration of physical, genetic and cytogenetic maps of human chromosome 7: isolation and analysis of yeast artificial chromosome clones for 117 mapped genetic markers. Hum. Mol. Genet. **3:** 489-501.

Green, E.D., V.V. Braden, R.S. Fulton, R. Lim, M.S. Ueltzen, D.C. Peluso, R.M. Mohr-Tidwell, J.R. Idol, L.M. Smith, I. Chumakov, D. Le Paslier, D. Cohen, T. Featherstone, and P. Green. 1995a. A human chromosome 7 yeast artificial chromosome (YAC) resource: construction, characterization, and screening. Genomics **25:** 170-183.

Green, E.D., D.R. Cox, and R.M. Myers. 1995b. The human genome project and its impact on the study of human disease. In The Metabolic and Molecular Bases of Inherited Disease. (eds. C.R. Scriver, A.L. Beaudet, W.S. Sly, and D. Valle), pp. 401-436. McGraw Hill, New York.

Gyapay, G., J. Morissette, A. Vignal, C. Dib, C. Fizames, P. Millasseau, S. Marc, G. Bernardi, M. Lathrop, and J. Weissenbach. 1994. The 1993-94 Genethon human genetic linkage map. Nature Genet. **7:** 246-249.

Heng, H.H.Q., X.-M. Shi, and L.-C. Tsui. 1993. Fluorescence in situ hybridization mapping of the cystic fibrosis transmembrane conductance regulator (CFTR) gene to 7q31.3. Cytogenet. Cell Genet. **62:** 108-109.

Hillier, L. and P. Green. 1991. OSP: A computer program for choosing PCR and DNA sequencing primers. PCR Methods Applic. **1:** 124-128.

Olson, M.V. 1993. The human genome project. Proc. Natl. Acad. Sci. USA **90:** 4338-4344.

Stapleton, P., A. Weith, P. Urbanek, Z. Kozmik, and M. Busslinger. 1993. Chromosomal localization of seven PAX genes and cloning of a novel family members, PAX-9. Nature Genet. **3:** 292-298.

Tamura, T., Y. Izumikawa, T. Kishino, H. Soejima, Y. Jinno, and N. Niikawa. 1994. Assignment of the human PAX4 gene to chromosome band 7q32 by fluorescence in situ hybridization. Cytogenet. Cell Genet. **66:** 132-134.

Walther, C., J.-L. Guenet, D. Simon, U. Deutsch, B. Jostes, M.D. Goulding, D. Plachov, R. Balling, and P. Gruss. 1991. Pax: a murine multigene family of paired box-containing genes. Genomics **11:** 424-434.

Zhang, Y., R. Proenca, M. Maffei, M. Barone, L. Leopold, and J.M. Friedman. 1994. Positional cloning of the mouse obese gene and its human homologue. Nature **372:** 425-432.

THIS PAGE WAS INTENTIONALLY LEFT BLANK

EXAMPLE: Human Ob Peptide is Biologically Active in Mice

Groups of 10 ob/ob mice were treated by i.p. injection with 10 μg/g/day recombinant (bacterial) human and murine ob peptide or saline. After four days, the group receiving saline gained 0.3 g. The group receiving murine ob lost 3.2 g. The group receiving human ob lost 2 g ($p < .01$ compared to saline controls). These groups were also tested for food intake. The data are shown in Table 5.

Table 5. Food intake/day (g) of treated ob/ob mice

Treatment	Day 0	Day 1	Day 2	Day 3
saline	13	13	12.9	13.2
murine ob	14	3	4	4.6
human ob	14.2	10.2	8.7	7.7

These data demonstrate that human ob is biologically active in mice.

EXAMPLE: A High Dose of Ob Affects Wild-type Mice

Wild type mice (C57Bl6J +/?) were treated with 10 μg/g/day i.p. of recombinant murine ob, and body mass measured every four days. The results are shown in Table 6.

Table 6. Body mass of normal mice receiving ob

Treatment	Day 0	Day 4	Day 8	Day 12	Day 16
saline	22 g	22 g	22.5 g	23 g	22.5 g
murine ob	22	20.5	20.7	20.8	21.8

These data demonstrate that ob affects the body mass of wild-type as well as obese (ob/ob) mice, albeit to a much smaller degree.

This invention may be embodied in other forms or carried out in other ways without departing from the spirit or essential characteristics thereof. The present disclosure is therefore to be considered as in all respects illustrative and not restrictive, the scope of the invention being indicated by the appended Claims, and all changes which come within the meaning and range of equivalency are intended to be embraced therein.

Various references are cited throughout this specification, each of which is incorporated herein by reference in its entirety.